

21 August 1980

Directive 59 and European defence

The Democratic convention last week was perhaps needlessly enlivened by the curious affair of Mr Edmund Muskie's claimed innocence of the most recent alleged shift in United States strategic policy — the increased emphasis now to be given to military targets in the disposition of American nuclear warheads. Several questions arise, but are unlikely to be answered in the near future, if at all. Did President Carter let slip news of the instruction to the US military on which the new strategy is based — called Directive 59 — in order to impress the impending convention or to take the wind out of Mr Reagan's sails, thus robbing the Secretary of Defense, Dr Harold Brown, of the place in the history books that he might have earned by making a public speech on the subject? (Dr Brown will indeed have made such a speech on 20 August.) Did Mr Muskie have to complain about not knowing of Directive 59 because his officials at the Department of State had not had time to tell him? And did he complain in the terms he used — more sorrow than anger — because of the rumours current earlier in the week that the Secretary rather than the President might end up as the Democratic Party's nominee for the election in November? In this quiet season of the year, these questions will excite attention for weeks to come. There is a danger that the more important question of what Directive 59 really means may be overlooked.

The first thing to say about Directive 59 is that it is not so much an order to the troops as an acknowledgement of what the troops have been doing, and are prepared to do. Many years have passed since the Department of Defence sought to demonstrate to Congress and the world that the great megatonnage in the warheads of Soviet missiles was more than made up for by the versatility of Multiply-Independent Re-entry Vehicles or independently targeted warheads, MIRVs for short. Since then, both sides of the equation have become larger, and their difference no easier to calculate. For at least a decade, however, the United States has made no secret of its wish to be so well equipped with accurate and separately manoeuvrable warheads that it can reasonably hope to attack crucial military targets, not just cities full of mere people. With the MX missile, and others of the same generation now in prospect, the US military (and their Soviet counterparts) should soon be able to deliver on that promise. Directive 59 is therefore largely an acknowledgement in public of what the Pentagon has been saying about its aspirations since the early 1970s. It should be acknowledged that half a decade has yet to pass before US Strategic Air Commands will be so well equipped that they can do what the US President now says they must do — to be prepared to fight strategic nuclear wars by matching their opponents' strength tit-for-tat while keeping enough in reserve to be able to deliver a final devastating assault. The commanders, however, are not running for election.

For those with long memories — historically, so to speak — there is an ironical circularity in this state of affairs. A quarter of a century ago, when the American Atlas and Titan rockets were still being tested, while the Soviet strategic armoury was dependent on subsonic aircraft for delivery, and when what was called the Cold War was at its height, a number of honest men (the late Sir John Slessor prominent among them) developed the theory of "graduated deterrence". The idea was that the world would be a safer place if the West, at least, was able to match its retaliation more or less exactly with whatever attacks might come its way. The objective was to find some way of managing the escalation of a strategic nuclear war so that it could stop short of the ultimate catastrophe, mutual and virtually total destruction of the two sides in a conflict. The concept was to many people preferable to

that of deterrence based on all-out retaliation at the first sign of a strategic attack, but even on paper there were snags. To many Europeans, the most obvious flaw in the theory was that controlled escalation of a strategic conflict would make it possible for the United States to lose its nerve half-way along, with the result that the United States guarantee of Western Europe would be diminished. The most serious argument against the theory of graduated deterrence, at least in the nuclear stages of a strategic conflict, was, however, technical; the missiles of the 1950s were simply not accurate enough to match the need. What has happened now, after a quarter of a century, is that it will soon be technically possible to follow a strategy of graduated deterrence.

What will be the consequences? Again the 1950s are instructive, and it is no surprise that Dr Harold Brown thought it wise to send a message of reassurance to the other NATO governments once Mr Carter had let Directive 59 out of the bag. The new strategic doctrine, Dr Brown said, is but a stage in the evolution of United States strategic policy. This is no more than he said at the meeting of the NATO Nuclear Planning Group at Bodo in Norway last June, and what he is likely to repeat when the same group meets again in November. Other NATO governments appear to have accepted this to be the truth, but they have not yet had time to work out the implications of the new development. The overriding question in their minds will be the question that preoccupied them in the 1950s — will the graduated deterrence strengthen or weaken the United States guarantee of Europe?

The calculation is far from easy, and is best accomplished by successive approximations starting from the assumption that both the American and Soviet strategic forces are capable only of the mass destruction of their opponents' cities. Although the North Atlantic Treaty, like the Warsaw Pact, is a solemn declaration of the superpowers' commitments to their allies, it is only a piece of paper. Nobody supposes that a strategic conflict in Europe would evoke an automatic response in Washington or Warsaw — people pressing buttons without turning over in their minds the prospect that their own people were about to be destroyed. In other words, even with the strategic balance of the 1950s and early 1960s, there was a chance that Europe might be left to its own devices and other people's missiles. The crucial question is whether Directive 59, old hat though it may be, increases or diminishes that chance.

First — this is what Dr Brown has been implying — it will be easier for the United States to honour its commitment to its allies if it can respond to a strategic threat by measures that are more specific than all-out retaliation. If, at the first sign of strategic trouble in Europe, somebody in Washington is faced with the choice of doing nothing or of setting in train an all-out nuclear exchange, he (or she?) is more likely to do nothing than if the choice were between inaction and some more calculated response. In this sense, Directive 59 should be a reassurance to the Western European members of NATO that their powerful ally will be more likely to stir a finger. By the same test, however, the controlled escalation of a nuclear conflict inevitably offers more opportunities for the United States (or, for that matter, the Soviet Union) to lose its resolution. Now, as was apparent in the 1950s, graduated deterrence cuts both ways. The strategic umbrella is broader, but there are more holes in it.

The immediate question (for when the US election has been held) is whether Western European governments will be more conscious of their increased security or of the increased opportunities for the United States to pull out from a European battlefield. Their perceptions will be determined by political and



not military calculations. Will a more bellicose president (Mr Reagan) also be more isolationist? Would President Carter, if re-elected, be able to balance his apparently incorrigible unawareness of the European consequences of his domestic pronouncements (such as that about Directive 59) with the flexible and imaginative view of Salt II, Salt III and the other matters at the root of European security to keep the confidence of his NATO allies? Time, no doubt, will tell.

Until that issue is resolved, there are bound to be hankerings in Western Europe after the position adopted by the French a quarter of a century ago, when the Gaullist doctrine of independence was evolving. The French position is quite clear, and has been consistent from the start. In the last resort, the argument goes, even the most solemn guarantees are likely to be abrogated. It makes no sense to think that the people of the United States will knowingly sacrifice themselves for the sake of France. On the other hand, France can safeguard its own future by acquiring an independent nuclear force able to inflict damage on potential adversaries that is more or less proportional to its own potential value as a prize. The British case for the retention of an independent deterrent, much rehearsed when the British government decided three weeks ago to adopt the Trident system of nuclear submarines as its next-generation nuclear deterrent, is subtly different. The objective, that argument runs, is to influence the perceptions not of allies but of adversaries: if it should be thought by others — however mistakenly — that the

United States might, in a tight corner, leave Europe to its own devices, the result might be a conflict that would otherwise be deterred. The West German government, denied both these options by its solemn declaration within the Western European Union that it would never become a nuclear power, is understandably (and rightly) more concerned with the threat from Soviet medium-range missiles across the Eastern frontier and would like faster progress with Salt III.

What, in these uncertain circumstances, is to be done? The conventional advice is to wait until after the American election (and then perhaps to wait until a new Administration is installed). That is a counsel of defeat. It is in the best interests of NATO, but would also serve the good cause of diminishing the risk of strategic conflict, if the three principal European states were to use the forthcoming period of uncertainty to decide among themselves what they want for European security. The British and French justifications of their independent nuclear forces are not so different that they could not be reconciled, while both are consistent with a NATO agreement with the Warsaw Pact on medium-range missiles (Salt III). Conventionally, again, it is supposed that an Anglo-French deterrent is unthinkable and that Salt III must wait on Salt II coming out of limbo (which in turn must wait on Afghanistan and, again, the election). But this is not a time for making decisions that must of necessity occupy most of the 1980s. President Carter's announcement of Directive 59, domestically, is an opportunity to begin that process.

What price the price of crude petroleum?

The recession in the industrialized countries of the West is now also beginning to hurt the oil-producing states. In the past few weeks, production in many OPEC states has declined as consumption in many industrialized states has stagnated or even fallen marginally. Some producers are now trimming back the premiums they have been in the habit of adding to the prices on which they have been settling at the periodic OPEC price fixings, with the result that the prices charged for refined products, especially petrol or gasoline, have themselves fallen by a per cent or so. Meanwhile, the Rotterdam spot market (which is not physically in Rotterdam but rather is the name for a system by which oil traders sell tankerfuls of oil to each other while the ships are still at sea) has seen the price of crude oil fall back and the volume offered for sale by the oil-producing states shrink to a trickle. Understandably, if prices are falling, the oil producers prefer to leave their petroleum in the ground. Only some of the oil producers, Iran most conspicuously, will be embarrassed if, for a time, oil revenues noticeably shrink. But the emerging slump in the demand for oil differs in many important ways from previous occasions in the recent past, 1976 for example, when there seemed to be a surplus of crude oil. It is also an opportunity for the oil-consuming states of the industrialized West to put their own pricing policies for oil in order.

There is some evidence that the present slackening of demand for oil reflects not merely the economic recession in the West but the beginning of an underlying trend towards reduced oil consumption. That this should be happening is merely a sign that increased prices bring lower consumption, the sort of thing that Adam Smith was saying 200 years ago. The only surprise is that it has taken seven years, since the first dramatic increases of the price of oil in the autumn of 1973, for consumers to make arrangements to manage with less. Ironically, in this the first stage of the recession, high interest rates will be a brake on the pace at which new capital developments might help to accelerate the conversion of industrial plant and commercial buildings to more efficient ways of using energy. It may be some small comfort for everybody concerned that, if this recession follows the standard pattern, there will come a time when interest rates are low again. On present form, however, even if present trends persist, it will be a decade or so before the industrialized West's dependence on imported OPEC oil has lessened to the point at which the

consumers will be able to look the suppliers in the eye and argue about the price of oil.

In the meantime, there are other steps that could and should be taken, the chief of which is to bring about a greater of harmony between the oil pricing policies of the consumer states. For governments appear still be divided on the objectives of their oil pricing policy: they see higher prices as a means of encouraging conservation, but they differ in the degree to which they wish to protect their electors from reality. The worst offender is the United States, where Congress has stolidly declined to let the Administration take steps to let the price of oil rise towards the international price of crude. (Is that another issue that will be more easily decided when the presidential election has been run?) One result is that the pressures towards conservation are not as strong as they might have been. Another is that energy-intensive industries in the United States are blessed with a commercial advantage over their competitors elsewhere which is widely considered to be unfair. In the past few weeks, there have been rumblings from the European chemical industry that European governments will not be able indefinitely to ignore. Yet they appear to be unwilling to follow the sensible course of harmonizing their systems for taxing oil so as to ensure that, at least within Europe, nobody has an unfair advantage. The European Commission's proposals to this end were given short shrift earlier this year. In Europe, perhaps the worst offender is the British government, which rigidly follows OPEC prices in trading North Sea oil without apparently appreciating the strategic benefits that would flow from a slightly lower price.

Further ahead, the overriding need is for a mechanism resembling a free market that would serve as a means of fixing a realistic price for oil and, ultimately, for other fuels. The notion that oil importers that are also oil producers (chiefly the United Kingdom and the United States) might deliberately use part of their production to increase the volume of crude oil traded on the Rotterdam market (see *Nature*, 3 July) could not fail to have some influence on future OPEC prices. The attempts now being made to organize a future market in crude petroleum, great though the technical difficulties may be, could be an influence of the same kind. The trouble, as always, is that Western oil producers are unwilling to sacrifice any part of their security of supply for the benefits (to them as well as others) of lower prices.

Columbia and Yale at odds about Felig

An unhappy scandal has rocked the departments of medicine at Columbia University and Yale University, has led to the resignation of the newly appointed chairman of the department of medicine at Columbia, Dr Philip Felig and has again raised accusations of plagiarism in the scientific literature.

The trouble, regarded as a tragedy for all concerned by members of the department of medicine at Columbia, was first reported by the *New York Times* last week. What follows is a chronology of the events of the past few years.

In December 1978, Dr Felig, then at the department of internal medicine at Yale, was sent by the *New England Journal of Medicine* a paper which had been submitted by Dr Jesse Roth from the National Institutes of Health with Dr Helena Wachslicht-Rodbard as co-author. The paper, entitled "Insulin binding in anorexia nervosa", was sent to Dr Felig as a referee and eventually returned to the *New England Journal of Medicine* where it was published in April 1979.

Early in 1979, Felig and his associate of five or six years, Dr Vijay Soman, submitted to the *American Journal of Medicine* a paper with essentially the same title as that of Roth, which was sent by the Editor to a group of referees which included Dr Roth.

The authors of the original paper recognized that some portions of the text of the Felig-Soman paper were identical with their own original and that an equation which became the basis for their own statistical analysis of their data was similarly identical. Dr Wachslicht-Rodbard protested to the Editor of the *American Journal of Medicine* and later Roth protested to Felig, as a result of which Felig offered to withdraw the abstract of the paper already submitted to that year's FASEB meeting in Atlantic City, to reprimand Dr Soman, to amend his paper so as to acknowledge Roth's priority and to arrange that the paper should not be published until after Roth's had appeared, which was in April 1979 (*New England Journal of Medicine* 300, 882).

Meanwhile, Roth's colleague at NIH had written to Felig and to the Dean of Medicine at Yale suggesting that much of the data which was the basis for the collaboration between Dr Soman and Dr Felig might be suspect and asking for an audit of the laboratory's work. Even though Dr Felig had at the outset denied the possibility that there might be irregularities in his laboratory, he agreed to an audit and, according to Dr Roth, not to publish his paper until the audit had been completed. In the event, the audit was long delayed and Soman and Felig's paper was published in January 1980 (*American Journal of Medicine* 68, 66). Only by

February 1980 had the independent audit been carried out, revealing discrepancies between raw data and published accounts thereof. Whereupon Dr Soman admitted "fudging" data.

During this period, a Columbia search committee was seeking a new chairman for the department of medicine. In the autumn of 1979, the search committee interviewed Dr Felig, decided to offer him the post and had completed negotiations with him for his removal from Yale to Columbia at the turn of the year. Since then, Dr Felig has been making periodic visits to Columbia, where he began work full-time at the beginning of the academic year in mid-June, having by then moved to New York.

People at Columbia say that they do not dispute Dr Felig's explanation that the alleged plagiarism in the article submitted to the *American Journal of Medicine* may have arisen because Dr Soman had kept a copy of the article sent to Dr Felig for review and passed on to him. Equally, they say, Dr Felig may have submitted his referee's report to the *New England Journal of Medicine* on the basis of a draft prepared by Dr Soman. They emphasize that there are no grounds for disputing Dr Felig's statement that he was unaware of similarities between his paper to the *American Journal of Medicine* and the paper submitted to the *New England Journal of Medicine*.

The audit in February 1980 of the data compiled in Dr Felig's laboratory prompted a second and again external audit by Dr Jerrold Olefsky of the University of Colorado. By then, Dr Felig had asked that Dr Soman should write a letter declaring that he alone was responsible for such similarities as there may have been between the Roth paper and that submitted by Felig and Soman to the *American Journal of Medicine* and Soman was dismissed. The independent audit disclosed that of more than a dozen papers submitted in recent years from Dr Felig's laboratory involving collaboration with Dr Soman at least eight were under a cloud and that six were unsupported by raw data gathered in the usual way in laboratory notebooks.

In February 1980, however, Dr Felig had proposed that Dr Soman should be appointed as assistant professor in the department of medicine at Columbia University without, people at Columbia now say, disclosing that Dr Soman's work was under a cloud. Although the dean of the medical school, Dr Donald F. Tapley, was informed of the outcome of the second independent audit of the Felig paper in March this year, it is now said that Dr Tapley did not inform his colleagues at Columbia until July, when word of the alleged plagiarism had been spread about and when Columbia, with Dr Felig already



Dr Philip Felig

in post as chairman of the department of medicine, set up an *ad hoc* committee to look into the affair. During the early summer of this year, Dr Felig was at pains to deal with the consequences of the two independent audits.

One of the papers concerned in the investigation — Soman, V. and Felig, P. "Regulation of the glucagon receptor by physiological hyperglucagonaemia" — was published in *Nature* (272, 829; 1978), and Dr Felig wrote to the Editor to say that it had not been possible to find the original data and that "until further data are available, the hormone binding data contained in that publication must be considered questionable". (Dr Felig had also written to the editors of several other journals in similar terms but a retraction in respect of another paper will appear in the current issue of the journal *Diabetes*.) Dr Felig was told by *Nature*, in May, that in Dr Soman's interest it would be best if a retraction in respect of the *Nature* paper appeared only when the position had been cleared up.

In the event, the *ad hoc* committee at Columbia took the view that Dr Felig was at fault in not having been able to recognize similarities between the paper which he signed and submitted to the *American Journal of Medicine* and that which he had, a few weeks earlier, been asked to review; for having delayed for as much as nine months before seeing that an audit of his laboratory's data had been carried out; and for not volunteering to his academic colleagues at Columbia the trouble he had encountered at Yale. Faced with this complaint, Dr Felig resigned from Columbia at the end of last month.

Yale takes a quite different view of what has happened. On the telephone this week,

Dr Robert W. Berliner said that Columbia had made a gigantic mistake. He said that the alleged plagiarism consisted of a sentence or two and an equation that is not nearly as special as is claimed. People at Yale, indeed, suspect that Dr Felig is the victim of a misunderstanding. He has been invited to return to Yale and has apparently agreed to do so. This week, Dr Felig was on vacation. Dr Roth, while sympathetic to the predicament in which Dr Felig initially found himself, expressed the view that it was careless of Dr Felig to go ahead with publication before the audit had been completed.

School computing

Money to spend

The British government's £9 million programme to promote the use of computers in schools is getting off to a slow start because of delays in appointing a director. To avoid the possibility of losing some of the money if it is not spent before next March, the Department of Education and Science (DES), which is putting up the funds, has already had to award some grants. About £500,000 of the £1 million set aside for this financial year has been either awarded or earmarked for projects which are as yet planned only in outline. The allocation of larger sums for projects which will define the emphasis of the four-year programme will have to wait for the new director whose appointment is expected to be announced in September.

The present programme has grown out of a plan of the previous Labour government to spend £12 million on developing the educational value of microcomputers. None of that money was spent, however, because of the change of government in April last year. After more than a year of deliberation, the present Conservative government announced last March that it was to renew the programme by spending £9 million over the next four years: £1 million is to be spent before March 1981 and £2.7 million is to be spent in each of the remaining three years. When it is fully underway, the programme is to be administered by the Council for Educational Technology for the United Kingdom.

The DES has yet to decide on the overall objectives of the programme because to do so it needs the advice of the director as well as a recently appointed advisory committee of civil servants, teachers and representatives of educational organizations. The advisory committee gave the matter some thought at its first meeting at the end of July. All members were agreed that the programme should aim to improve the use of computers in schools but there was a difference of opinion on what should be the long-term objectives of computer education. Some members believe that the aim should be to acquaint most school leavers with computers and their uses:

others believe that the aim should be to give pupils a grounding in programming and computer technology. Members plan to resolve the issue at the committee's next meeting in September.

Much of the work on computers in schools has been done by the Council for Educational Technology for the United Kingdom, which will administer the current programme when it gets under way. Other groups, which will no doubt be looking for funds, include Micro and Mini Computer Users in Secondary Education, a voluntary organization which is building up a library of software and guidance notes for teachers, and the Schools Council. The Schools Council hopes to receive grants to extend its existing work on compiling software and course notes on several subjects central to the school curriculum. Some of the packages it has already produced are designed to be of use in teaching history, economics, geography and even home heating, as well as chemistry, physics and biology.

Only about one hundred schools in Britain own their own computer and it is unlikely that much of the £9 million will be spent directly on increasing that number. One major constraint on the work funded under the programme will be that it should lead to results or suggestions for computer use which are within the financial limits of local education authorities.

Judy Redfearn

Soviet environment

Old hands back

Environmental protection is now a major feature of Soviet planning, and Moscow radio recently took to task Western newspapers (including the *Guardian*) which, it said, reported only the negative aspects of Soviet ecology. In response, the commentator drew specific attention to a speech of Academician N.P. Dubinin, delivered at this year's plenary meeting of the Soviet Academy of Sciences, and which was later published in the *Vestnik* of the Academy.

Academician Dubinin is a leading Soviet geneticist, whose work, notably his book *Evolutionary Genetics* had been

suppressed during the Lysenko period. His address to the academy, entitled *Genetics and its significance for Mankind* covered a whole range of topics from plant-breeding to genetic engineering. Nearly half his speech, however, was devoted specifically to the genetic aspects of pollution, which, he explained, make up an important section of the Soviet Union's "Man and biosphere" programme, which is the responsibility of a special environmental Interdisciplinary Council of the State Committee of the USSR on science and technology.

In his speech, Dubinin stressed that much of the work in this field was taking place within the framework of the United States - USSR collaboration programme. By stressing the collaborative aspect, he was thus able to quote exclusively from US sources on the genetic hazards of pollution by mutagens, or to attribute to American authors the more pessimistic viewpoint that "at the present time science does not have at its disposal the means of solving this exceptionally complex problem". On the other hand he attributes to Soviet authors a "highly promising express test scheme" for studying the frequency of genetic mutation, tests based on leukocyte cultures and a method for determining gene and chromosome mutations in embryo mouse cells. The American J. V. Neel, however, receives due credit for his improved monitoring programme for observing the increase of mutations due to the Hiroshima and Nagasaki bombs.

In a speech intended to review the positive aspects of Soviet genetics, open criticism of past policy would clearly be out of place. There were, however, one or two, interesting allusions to the Lysenko aberration. In speaking of the "adaptive norm", Dubinin specifically cited his own (erstwhile banned) work of 1948, and stressed the importance of establishing whether or not unfavourable genetic changes had occurred over the last 20 to 30 years. Soviet genetics, it should be recalled was effectively restored only in 1963 - 64. In a speech which the Soviet media singled out as the proper and positive way to report environmental genetics, such low-key criticisms, it appears, are quite in order.

Vera Rich

Frogs to eat



The edible frog (*Rana esculenta*), the basis of a number of French recipes for gourmets, is one of the species widespread on the mainland of Europe which have established themselves only with difficulty in the United Kingdom. The photograph is taken from a Nature Conservancy Council booklet, *Wildlife Introductions to Great Britain*, which is the report of a working party of the UK Committee for International Nature Conservation. The edible frog is apparently now established only in the counties of Norfolk and Sussex.

Interferon production

Japan now

International agreements for the manufacture of interferon appear to be catching on. Last week, the Wellcome Foundation Limited (whose sole shareholder is the British foundation known as the Wellcome Trust) announced an agreement with the Sumitomo Chemical Company of Japan that will in due course enable the Japanese company to use Wellcome technology for manufacturing lymphoblastoid interferon.

The project, which is thought to be potentially worth several million pounds to the British company, will first of all involve clinical trials in Japan based on interferon supplied by Wellcome. Thereafter, the Japanese company will decide whether or not to go ahead with the building of a plant for producing interferon commercially, using the techniques developed at the Wellcome Foundation over the past few years.

Broadly speaking, the technique is to infect a culture of lymphoblastoid cells with Sendai virus, which stimulates interferon production, and to extract lymphoblastoid interferon on chromatography columns loaded with antibodies against the type of interferon being produced.

So far there has been no opportunity to decide whether the recently announced technique for raising monoclonal antibodies against interferon (Secher, D. S. and Burke, D. C., *Nature*, **285**, 446; 1980) will make this process efficient.

The cost of interferon plants of this kind is likely to be of the order of £5 million. It is thought that the production of interferon by the Wellcome Foundation now exceeds the total production of Finland, hitherto the chief source of interferon. For those drug houses manufacturing interferon from tissue cultures, the overriding question for the future must be whether cloning techniques will lead to more efficient and cheaper manufacture.

If it is possible to obtain as many as a hundred molecules of interferon from a bacterium such as *Escherichia coli* carrying an interferon gene, the capital costs of the manufacturing plants required to produce a given amount of interferon will be reduced at least in proportion to the ratio of the generation times of bacteria and cells in culture — 20 minutes to three days, or a factor of a hundred. Drug houses using tissue culture techniques (which are thought to include Roche and Merck as well as Wellcome) appear however to be counting on the length of the period of clinical investigation of cloned interferon that would be needed before these could be put on the market.

As things are, there is no direct evidence that cloned interferon is equivalent to that from tissue culture and in any case the multiplicity of interferon molecules now

being identified will need to be distinguished from each other physiologically and clinically, a process thought by the existing manufacturers to require five years or so.

Although the agreement with the Sumitomo Chemical Company has been formally signed, there are some grounds for anxiety that the arrangement may be frustrated by difficulties in Japan about the widespread use of whatever materials the Wellcome Foundation's Japanese partner eventually seeks to put on the market.

Occupational safety

Cuts ahead

Britain's Health and Safety Executive (HSE) is expecting difficulty in cutting its coat according to its cloth. In its latest annual report, published earlier this week, the HSE says that the government's plan to set its 1981-83 grant at 6% below the grant for 1979-80 will mean a reduction in the amount of work it can carry out.

Most of the savings will be made on staff, probably by reducing numbers from 4200 to 3800, and on internal office costs. But, says the report, some of its work of direct relevance to employers and workers, in particular that of the Factory Inspectorate, is bound to suffer. The research budget is also expected to come off particularly badly, with cuts of 18 per cent or about £2.2 million expected for the main research and development budget.

The main fear of John Locke, Director of the HSE, is that its work will become more reactive, dealing only with problems retrospectively. Such a move would be contrary to the policy of anticipating problems adopted by the executive since its creation in 1974.

Locke regards the anticipation of problems through extensive visits to workplaces and an active research programme a priority, even above some of the daily demands on the executive's time. In the report, he says that longer term work on problems with no immediate obvious effects on employers and workers, especially those of occupational exposure to chemicals, noise and radiation, must continue even if cuts have to be made elsewhere. Such a policy would be in line with the changing emphasis of the executive's work over recent years from accidents at work to longer term health problems.

The plan is to make most of the cuts in the research budget fall in those areas, commonly involving standard testing, where industry could be persuaded to step in with funds. Hence, the HSE hopes that industry will make up for the £1.6 million reduction in the research division's direct support to the inspectorates by arranging for its own tests to be done. Industry is also to be encouraged to pay for and conduct some of the work done under three of the executive's major current research projects;

a study on the toxic effects of low levels of isocyanates, a study on mortality in the rubber industry and a project at Porton Down to study the dispersal of heavy gas clouds. To avoid industry totally controlling studies on its own problems, the HSE still intends to be responsible for designing and monitoring research projects. Industry, however, will be encouraged to provide its own data for them or to pay for research that is done independently.

One of the executive's problems in defining a coherent programme of future work is that with the increasing public awareness of health and safety issues, largely fostered by its previous work, increasing demands are constantly being made on its resources and new problems are brought to its attention. One of the major areas in which it will be stepping up its work is nuclear safety to cope with the government's plans to increase the nuclear power programme and to build a pressurized water reactor. Additional resources are to be put into the Nuclear Installations Inspectorate and almost all the work on fast reactor safety is to cease so that resources can be transferred to studies of PWR design. Work will also increase on toxic chemicals and major hazards, both of which are part of EEC programmes of research.

In a plea for kinder cuts by the government, the report argues that a major reduction in its services could lead to increased expenditure elsewhere, especially in the health services and sickness pay. It says that if further cuts, beyond 6% are planned then the Health and Safety Commission, the body to which the executive is responsible, would look to the government for suggestions as to which areas of its work it should reduce.

The executive does have some statistics to support its case. The number of fatal accidents for those industries covered by the executive throughout its five year life, fell from 651 in 1974 to 544 in 1979. Non-fatal accidents have also dropped by about 25% over the period.

Judy Redfearn

Graduate students

Places sought

Demand from new British graduates for research grants is surprisingly buoyant this year. Half-way through the short season during which the research councils make awards, the Science Research Council at least is modestly encouraged by the volume of applications, which is well up on that of last year.

Part of the explanation is the relaxation of the rules for the awards in the gift of the SRC Engineering Board. For the second year, awards of engineering studentships both for research and for advanced study (usually leading to an MSc degree) are being made on the principle of "first come,

first served'.

Now that word of the new regulations has percolated to the universities, the Engineering Board's objective of increasing the amount of postgraduate study seems well on the way to being achieved. The Engineering Board has set itself a target for this year of 435 research studentships and 535 advanced course studentships, and gives the impression of being prepared to spend more in future years if only it can attract suitable applicants.

In science subjects, where the rules are unchanged, the target for the award-making season is virtually the same as last year — 769 research studentships and 435 advanced course studentships. These awards are made on the nomination of university and polytechnic departments on the basis of quota allocations by the various subject committees of the Science Research Council. Even so, there is some evidence that the volume of applications is greater than in 1979.

The parallel operation of these two systems is likely to throw an interesting light on the workings of the quota system. Although (outside engineering) the quotas are fixed by the SRC subject committees after a necessarily subjective estimate of the capacity of individual departments to provide training in research, many of those who wish to see a further concentration of research support on departments with high reputations have regarded the quota system as an impediment to change. Would-be engineering research students are in the circumstances more free to decide where to settle, for which reason a comparison of the distribution of research students in science and engineering will be of some interest when the award-making process is over.

Applications for postgraduate awards have also increased under the SRC's scheme for Cooperative Awards for Science and Engineering (CASE) although a relaxation of the rules may again have helped. CASE studentships involve research projects worked out in partnership by a university department and an industrial company, which in previous years have had to be approved by the SRC before a student award can be made. For the first time this year, university departments have been allowed to nominate a student while seeking approval for a project, although an award remains conditional on the students' performance in the degree examinations.

The target for CASE studentships this year is 760, but clearly the SRC would again spend more if there were a sustained demand. It is, however, something of a surprise that the council has been able to build up to 700 awards a year in less than a decade of CASE studentships.

The apparent buoyancy of demand for postgraduate awards is somewhat at odds with reports that many new graduates are this year more anxious to find permanent

jobs than to remain in research. By the end of the summer, it is also possible that much of the demand may have melted away. One of the hazards of the process is that awards offered are not always taken up, sometimes because applications are made in duplicate. The SRC's record of this kind is that of a husband and wife who each made sixteen separate applications.

NPT review conference

Smooth start

The Second Review Conference of the Non-Proliferation Treaty (NPT) got off to a smooth and predictable start in Geneva last week. From the opening declarations which delegates to the conference were making in the first six working days, it is plain that the non-nuclear weapons states which are parties to the treaty will be urging that the nuclear powers have made too little progress too slowly towards strategic disarmament, and that the restrictions imposed on trade in nuclear materials in recent years are both a violation of the treaty and discriminatory against developing countries.

Delegates from nuclear weapons states are however heartened that these protests have been couched in moderate language. For their part, the nuclear weapons states (the Soviet Union, the United Kingdom and the United States) have been making much of the report on the comprehensive test-ban negotiations submitted to the Committee on Disarmament at the end of July and of the Committee on Assurance of Supply set up under the International Atomic Energy Agency and due to hold its first meeting at the end of September.

Procedural matters appear to have caused so far the few anxieties that have flitted across delegates' minds at Geneva. As planned by the preparatory conferences, however, the Iraqi delegate was in the end appointed chairman of the conference. It has been agreed that most of the work will be done by two committees which met for the first time on Tuesday and which are concerned respectively with progress towards disarmament and security of supply.

Work on the drafting of the conference report will begin next week, and although everybody seems reconciled to the notion that the report will refer to the discontents of the non-nuclear powers, it was hoped last week that a consensus document would be agreed.

For the most part the nuclear powers presented a united front in their formal speeches, drawing attention to the need to strengthen the NPT system and to the detailed character of the report on the comprehensive test-ban negotiations. At one point, however, the Soviet delegate did seek to embarrass the United States by pointing out that Salt II would by now have been ratified had it not been for the United States Congress.

The discontents of the non-nuclear powers, expressed in several of the formal statements, were eloquently put by Mr Domingo Siazon, the leader of the Philippine delegation, who argued forcefully that restrictions on the supply of nuclear materials devised in the past few years by nuclear suppliers were a violation of the NPT, and that signatories of the treaty such as the Philippines, having signed in good faith, had every reason to expect that the nuclear powers would honour their undertaking to provide nuclear materials, fuel and equipment, without discrimination.

It is planned that the review conference will end on 5 September.

Windmill power

UK gets wind

Windmills seem to be on the march in Britain. Officials at the Central Electricity Generating Board (CEGB) will embark this summer on the choice of a site for a large windmill which the board plans to order in 1983 and commission by 1985.

The decision to buy a windmill, announced last week, seems to indicate that the board has decided to take wind power seriously. It is the first piece of alternative energy technology to be bought by the board, Britain's sole electricity utility, on a commercial basis. The early announcement of the plan is intended to stimulate British industry to develop an improved machine by 1983.

The CEGB's plan is to choose a site where other windmills could be built later to assess the characteristics of windmill arrays. It is prepared to spend about £10 million on building ten 1-MW machines. The first windmill will be a 1-MW machine, probably of horizontal axis, with a blade span of no less than 200 ft suspended 150 ft above ground. While industry is busy working on new designs to compete for the order, the CEGB will be gaining operating experience on a smaller 100-kW machine which it hopes to buy as soon as possible.

The apparent sudden interest in wind power has been stimulated by recent developments in the United States. Hamilton Standard, an American company which has pioneered wind-turbine technology, has recently claimed to have designed a windmill which could bring down the cost of wind-generated electricity to as little as 1–2 pence per kWh. Previous best estimates of cost have been about 3–4 pence per kWh.

Undoubtedly, these low estimated costs have stimulated interest in wind power within the CEGB. It is too soon to know what part windmills will eventually play in the pattern of the CEGB's electricity generation. Intermittent generation is an obvious problem unless pumped storage is available. The CEGB says that the chief contribution will be in saving fuel.

Costs of generating electricity from

power stations now being built range from 2.28 pence per kWh (nuclear) to 3.59 pence per kWh (coal). If, because of intermittent generation, windmills do not reduce the total need for installed capacity, however, the cost of wind-generated power should be compared with the costs of fuel in conventional stations, 0.61 pence per kWh (nuclear) and 2.30 pence per kWh (coal).

Hamilton Standard claims to have applied some of the principles involved in building helicopters to the design of a flexible windmill capable of bending in the

wind and responding to wind coming from different directions. Until now, most windmill designs have assumed the need for a large rigid structure capable of standing up in gales and for very strong rigid sails to withstand opposing gravitational forces. Hamilton Standard is now saying that a bendy windmill can withstand considerable forces and can be built more cheaply because less material is needed.

A cheaper windmill could also operate economically in wind speeds of only 6 m s^{-1} to operate effectively. The ability to

respond more sensitively to variations in wind direction also suggests that a flexible design of windmill might be successful in relatively low wind speeds.

In the light of these considerations, the CEGB says that it will be looking at flat inland sites, probably in the east of England, as well as hilltop sites for its first big windmills. A low inland site would have the added advantage of causing less environmental trouble than a hilltop in the rural areas of Scotland or Wales.

Judy Redfearn

Nuclear generation costs compared

The report of the Central Electricity Generating Board (CEGB) for the year to the end of March 1980 includes a detailed analysis of the costs of generating electricity from power stations of different types which is likely to be much quoted in the months ahead, when plans are being laid for the further development of the British nuclear power programme. The figures now published represent a departure from previous practice in that the assumptions on which they are based are stated more explicitly than in the past, with allowance being made for elements of cost such as the interest paid during the period of construction and the cost of decommissioning the stations concerned.

During the past year, the cost of nuclear power generated from the first generation of British "Magnox" stations was less than that of electricity from coal or oil. The costs for the past year are shown in the accompanying Table 1.

Part of the reason why the nuclear costs turn out to be so advantageous is that the initial capital costs of the Magnox stations, built between the late 1950s and the late 1960s, are much lower than present-day replacement costs would be. Although the same is true of the capital costs of conventional power stations, the fact that nuclear capital costs represent a greater proportion of the total implies that the figures are not a reliable guide to the costs of building new power stations of the same types.

One surprising feature of the CEGB's operations during 1979–80 is that the load factors of the nuclear stations were more or less the same as those of conventional stations. Electricity generated by the nuclear plant amounted to 70 per cent

of net capacity, compared with 69 per cent for coal-fired stations and 77 per cent for oil-fired stations. Originally, it was intended that the nuclear plant in the British generating system should be run at virtually full capacity; the board says that the lower load factors now being obtained from the nuclear power stations reflect their increased age. The costs of electricity generation are obtained by dividing the total cost of operation by the amount of electricity generated and thus increase as load factors decline.

Table 2 Costs for Hinkley B (AGR) and Drax (coal-fired)

	Hinkley Point B p/kW h	Drax first half p/kW h
Capital charges and provision for decommissioning	0.37	0.12
Interest during construction	0.18	0.04
Inclusive fuel costs	0.55	1.25
Other costs of operation	0.16	0.09
Research	0.07	0.01
Training	0.02	0.01
Total	1.35	1.52

The CEGB's annual report also provides costs on a comparable basis for two of its newest power stations — the Advanced Gas-Cooled Reactor station (the first of its type to be commissioned) and the newly-completed Drax coal-fired power station. The figures are shown in Table 2. The load factors for the AGR station (Hinkley Point B) and Drax were 55 per cent and 74 per cent respectively. The poor availability of the AGR reflects the difficulties in bringing this station into operation. It may seem remarkable that, in the circumstances, the cost of electricity generated is not much greater than that from the older Magnox station.

Mainly because of inflation, the CEGB

expects nuclear costs to rise. The net cost of electricity from the three AGR stations soon to be commissioned (Dungeness B, Hartlepool and Heysham I) are estimated at 2.62 pence per kWh, 2.28 pence per kWh and 2.31 pence per kWh respectively. The increase is largely attributable to increased capital costs, three times as great as those for Hinkley Point B. At the same time, the board estimates that the cost of conventional electricity will have increased still further, with oil-fired plant producing electricity at a cost of more than 6 pence per kWh and the second half of the coal-fired Drax station producing electricity at 3.59 pence per kWh.

For the first time, the CEGB also provides figures to show the basis on which future investment decisions are being made. The principle is that of discounted cash flow in which future costs are represented by a present cost using a discount rate of 5 per cent per annum. The board also makes the assumption that the price of coal will increase at an average rate of 2 per cent a year at least until the end of the century and notes that the cost of building generating plants of all

Table 3 Overall comparison of costs for nuclear and coal-fired stations

	Nuclear p/kW h	Coal-fired p/kW h
Capital charges at station and provision for decommissioning*	1.39	0.76
Interest during construction*		
Inclusive fuel costs	0.61	2.38
Other costs of operation	0.22	0.21
Generation costs	2.22	3.35
Less fuel saving from displacing less efficient plant	2.68 —0.46	3.02 +0.33

kinds has recently been increasing faster than inflation. The calculations now published, and summarized in Table 3, show that a nuclear generating plant has potential advantages over a coal-fired plant both in generating electricity more cheaply and in its effect on the economics of the generating systems as a whole — assuming that nuclear plants would displace less efficient plants, the Net Effective Cost of nuclear generating capacity becomes negative, implying a net saving for each power station built.

Table 1 Costs of power generation by Magnox, coal-fired and oil-fired stations in 1979–80

	Nuclear (Magnox) p/kW h	Coal- fired p/kW h	Oil- fired p/kW h
Capital charges and provision for decommissioning	0.34	0.09	0.14
Interest during construction	0.06	0.02	0.02
Inclusive fuel costs	0.60	1.29	1.61
Other costs of operation	0.26	0.14	0.14
Research	0.03	0.01	0.01
Training	0.01	0.01	0.01
Total	1.30	1.56	1.93

CORRESPONDENCE

Nuclear facts

SIR — The letter which you published on 10th July under the heading "Nuclear Protest" from three correspondents from Edinburgh Science for People, contains several misleading statements.

It is wrong to claim that trade union rights are "denied" in the nuclear power industry. Many of the staff in the industry who belong to trades union are very active in participating in debates about nuclear issues within their movement and in promoting the case for nuclear power development. They are presumably best able to judge their own conditions of employment; your correspondents should take account of their views and of those of the trade union leaders who speak for them.

It is incorrect to describe the UKAEA Constabulary as "an armed secret police force". There is nothing secret about them. They are sworn in as special constables. They are individually responsible for their actions and their powers are more circumscribed than those of the ordinary civil police. The AEA Constabulary are responsible to the Secretary of State for Energy through the Chairman of the Authority, and their functions have been reported to and debated by Parliament.

It is impossible to claim that the civil application of nuclear energy to commercial electricity generation has led to weapons proliferation; all of the existing nuclear weapons states found ways of implementing their political decision to have such weapons which did not start from a nuclear power programme.

This exchange of letters seems to me to substantiate the need identified in your editorial (19 June) for professional people, in universities, industry and government, to resolve their points of disagreement. We should at least be able to start with a correct statement of the facts before debating the political aspects of nuclear power development.

L. E. J. ROBERTS

AERE Harwell,
Oxon, UK

Geological sport

SIR — The Geological Olympics are over for another four years. The 26th International Geological Congress held on 7–17 July in Paris had the feel of an Olympic Marathon, with well over 2,500 talks in nine working days. Anybody attending each day deserved an Olympic medal for endurance, for on some days they had to make an agonizing choice between eighteen simultaneous sessions on such diverse topics as Ophiolites, Cambrian Stratigraphy, Geophysics of the Lithosphere and Upper Mantle, Applied Geochemistry, History of Tectonics pre 1922, Coal and Magnetic Stratigraphy and Ocean Floor Reconstruction.

What was the delegate likely to encounter? There was always the chance of "meeting old friends and hearing old papers" but in fact there was plenty of new research "on display". The problem was the way in which it was displayed. Having sat through approximately 2,000 slides (16 papers/day $\times$ 6.5 days $\times$ 20 slides/paper = 2,080 slides), I can honestly say that probably only one half

of them were legible. The rest might well have not been shown — indeed, some speakers even apologized with the standard "If you could see the relevant points, you would notice. . .".

Why are some slides so atrocious at conferences? Admittedly some were superb — large lettering, good use of colours. Others, regrettably, were just unintelligible. The slide of a table photographed directly from a journal, itself produced by offset-litho techniques is a good example. In these cases, the printing has probably been reproduced three or four times with an inevitable reduction in definition each time. The result is a blur from row 1, a sigh from the back and another author loses his audience's concentration. Maybe the time has come for "pre-talk screening" of slides to take place and rejection of those considered of dubious quality.

Apart from poor visual aids, the conference was a tremendous success. The sight of two German scientists actively discussing manganese nodules in English at a conference in France under a session chairman from Hawaii confirms that the occasion had an international flavour. But I wonder how many tribal groups (geochemists, geophysicists, petrologists, etc.) actually strayed outside their own compounds during the fortnight and how many merely "flew in" for the relevant day's papers.

Those who took this second course will have undoubtedly missed some of the more worthwhile aspects of the gathering — although no doubt pressing commitments elsewhere are partly to blame for brief visits. Maybe the Earth Scientists of today are becoming too specialized and hence, as chemists become geochemists, physicists geophysicists etc., the ability to understand broad based geology is declining.

S.J. WAKEFIELD

Dept. of Oceanography,
University College,
Swansea, UK

2,4,5-T dispute

SIR — Now that I have returned to the UK I have seen your editorial of 17 July and without complaining about the snide reference to me (no doubt prompted by your reluctance to publish the piece you requested me to write on the cancer issues) you have still got the 2,4,5-T issue wrong. I have no idea where you got the figure of 100 tons for the quantity of herbicide-based materials imported. As far as I know it is nearer 160 tons. *Nature* should (as in the back part of the magazine) abandon prejudice and support quite reasonable and sane crusades to protect people.

CLIVE JENKINS

Association of Scientific Technical
and Managerial Staffs, London, UK

The source of the figures quoted in the leading article to which Mr Jenkins refers is as follows. Customs and Excise give imports of 2,4,5-T into the United Kingdom in 1979 as 116 tonnes. Use of the material in the United Kingdom in 1979 was given as 58 tonnes in an answer to a Parliamentary Question on 20 June. The difference is accounted for by re-exports. Mr Jenkins has been informed of the source of the figures used but insists that his letter should be published. The article of his own to which Mr Jenkins refers was invited in March but not published because, in the judgement of the editorial staff of *Nature*, it dealt with a different aspect of the problem from that originally defined.

Editor, *Nature*

Misciting latest

SIR — In their News and Views article "Just so stories and cautionary tales" (*Nature*, 31 July)¹, Robert May and Miranda Robertson quote our study² where we point out that W. D. Hamilton's paper *The genetical evolution of social behaviour*³ is frequently miscited as *The genetical theory of social behaviour*. Although the recent trend to miscite the paper can be traced back to E. O. Wilson's *Sociobiology, the New Synthesis*⁴, the mutation has occurred independently at least three times and is an easy one to make for those familiar with Sir Ronald Fisher's *The Genetical Theory of Natural Selection*⁵.

Similarly, misciting Wilson's book as *Sociobiology, the Modern Synthesis*, as May and Robertson did, is easy to do for anyone who has read Sir Julian Huxley's *Evolution, the Modern Synthesis*⁶.

PAUL HARVEY
JON SEGER

School of Biological Sciences University of
Sussex, Falmer, UK and Department of
Biology, Harvard University, Cambridge,
Massachusetts

1. *Nature* 286, 327–329 (1980).
2. *New Scientist* 87, 50–51 (1980).
3. *J. theor. Biol.* 7, 1–52 (1964).
4. Harvard University Press (1975).
5. Clarendon, Oxford (1930).
6. Harper, New York (1942).

Distance measured

SIR — Bougault and Treuil in their *Nature* article (17 July) fail to inform us whether their distances are nautical or statute "mille passum" — there is a difference of 11.5 chains between these units, equivalent to 10,000 yr of seafloor spreading. Confusion could be avoided, and precision improved, by quoting distances in the traditional leagues of Columbus, Champlain and the imperturbable Captain Nemo.

PAUL MOHR

Department of Geology,
University College Galway,
Galway, Ireland

Filter tips

SIR — Dr Zeki's red filter described in your article (*Nature*, 31 July, p.435) is really excellent for use in photographing fluorescent DNA bands on agarose gels. One of the most important factors for a contrast filter is the absence of fluorescence of the filter itself. Conventional red filters usually emit fairly strong fluorescence when hit by a strong UV ray from a transilluminator, and this fluorescence makes for a somewhat bright background and an unclear photograph. We have been using a UV filter in front of a conventional red filter, to cut UV light and thus to minimize the self-fluorescence of the red filter. In this manner a satisfactory photograph can be obtained. It is important to use a good UV filter, however, because some emit a weak fluorescence.

Comparison of Dr Zeki's filter with our double filter showed the former to be superior. Dr Zeki's filter is completely free of intrinsic fluorescence. My thanks to Dr Owen for his helpful observation.

TAKESHI UOZUMI
Department of Agricultural Chemistry,
The University of Tokyo, Japan

NEWS AND VIEWS

Neutrino weight watching

from A. De Rújula and S.L. Glashow

In the past few months, there has been a great deal of excitement about neutrinos, chiefly because of two sets of experiments which have been reported. One experiment, due to F.W. Reines and his group at the University of California (Irvine), based on measurements of neutrinos from a nuclear reactor, has been interpreted as evidence for neutrinos whose mass is not identical with zero. The same conclusion is suggested by very preliminary results from a neutrino experiment at CERN. To the extent that both these experiments suggest that there may be interconversion between the three known species of neutrinos, the conclusions would help to account for the measured deficit of neutrinos from the sun. The interpretation of the experiments so far reported is however both uncertain and controversial. Plainly we are in for a period of lively argument; the hope must be that people will now design the experiments that must and should be carried out to settle this intriguing issue.

Pauli, in 1932, postulated the existence of neutrinos (and antineutrinos) in order to explain why beta rays, unlike alpha and gamma rays, are not monoenergetic. Nuclear beta decay releases a fixed amount of energy shared between two light particles, an electron and an antineutrino (so that $n \rightarrow p + e^- + \bar{\nu}$). Pauli's particles were not however 'seen' until 1957, when Cowan and Reines showed that antineutrinos emitted by a nuclear reactor produce positrons in a remote detector ($\bar{\nu} + p \rightarrow e^+ + n$).

The muon (μ) is a charged particle much like the electron, but heavier. It is produced together with a neutrino in the decay of a pion by a process much like nuclear beta

decay. In 1962, we learned that neutrinos produced with muons (ν_μ) are distinct from those produced with electrons (ν_e) for a beam of 'muon neutrinos' arising from decaying pions was seen to produce muons in a detector, but never electrons. Today, the electron is known to have an even heavier charged sibling called tau (τ). It, too, seems to have its own kind of neutrino (ν_τ).

Experimental study of tritium beta decay a decade ago demonstrated that the mass of the electron neutrino may be zero and must be very small, less than $60 \text{ eV}/c^2$ or 10^{-31} grams¹. The mass is at least four orders of magnitude less than that of the electron. Perhaps all three kinds of neutrinos are exactly massless — what could be simpler and more elegant? The trouble is that this kind of reasoning once led us astray with space reflection, time inversion and charge conjugation, which have turned out not to be symmetries at all. Nature will find her own way to elegance. We dare not frame laws of nature that have no logical necessity. Conservation of energy and of electric charge are well founded. Many other 'laws' are suspect. Protons may not live forever, and neutrino masses may be merely very small.

Indeed a recent report from the Institute for Theoretical and Experimental Physics in Moscow suggests that the mass of the electron neutrino is not zero². Again, the beta decay of tritium was studied, because it is the least energetic of all beta decays and hence the one most sensitive to neutrino masses. According to the authors, "the results perhaps give an evidence for the existence of the non-zero neutrino mass" in the range 14 to $40 \text{ eV}/c^2$. This cautious conclusion reflects the great difficulty of the analysis of the experimental results, for the effects of such neutrino masses on the observed energy of the electrons are of the same order of magnitude as the atomic or molecular corrections that must be made. Astrophysicists must nevertheless be delighted by the Soviet result, since they have a problem that neutrino masses of a few eV or tens of eV may solve. There is more mass in our universe, it seems³, than

can be accounted for in known forms of matter. This 'missing mass problem' recurs at various scales of length — the universe as a whole; in the dynamics of clusters, groups and pairs of galaxies; and in the massive, extended and non-luminous halos that most galaxies are known to have. May this missing mass be in the form of neutrinos?

Cosmological black-body radiation is interpreted as a relic of and evidence for the primordial big bang. As well as the observed electromagnetic radiation, it must include neutrinos. Hundreds of neutrinos, on the average, must inhabit each cubic centimeter of the universe. Should the heaviest neutrino have a mass of between 5 to $50 \text{ eV}/c^2$, these relic neutrinos would *in toto* outweigh the stars. They would provide the missing mass, and they could bind to galaxies to form their halos. The mass in our universe would consist mostly of neutrinos, which are so weakly interacting that they may be detected only by their collective gravitational effects.

Non-zero neutrino masses imply a new phenomenon called neutrino oscillation, or neutrino mixing. Bruno Pontecorvo⁴, in 1968, showed that massive neutrinos might change their identities in flight. A particle which is born as an electron neutrino in a beta decay may periodically behave as if it were a muon neutrino or a tau neutrino. Suppose that there are three kinds of neutrinos and that they are not massless. Call them ν_H , ν_M , and ν_L : the heaviest, the middleweight, and the lightweight. The crucial point is that ν_e need not be one of *these* neutrinos. Its wave function may be a linear combination of the three, with the muon neutrino and the tau neutrino being orthogonal combinations of the same three. On this picture, neutrinos arising from nuclear beta decay are born as electron neutrinos but, as time passes, their ν_H , ν_M , and ν_L components evolve differently. As the linear combination changes, the beam becomes a mixture of electron neutrinos, muon neutrinos and tau neutrinos. There will be oscillations of neutrino identity with frequencies that are determined by the differences between the masses of the neutrinos. These oscillations can lead to a variety of novel observable effects.

Electron neutrinos should be copiously

1. Bergkvist, K.E. *Proc. Topical Conf. Weak Interactions* CERN 69-7 p.91 (1969).
2. Lyubimov, V.A., Novikov, E.G., Nozik, V.Z., Tretyakov, E.F., & Kosik, V.S. (unpublished).
3. For a review, see: Faber, S.M. & Gallagher, J.S. *Ann. Rev. Astron. Astrophys.* 17, 135 (1979). A more recent discussion is found in Schramm, R. & Steigman, G., Gravity Foundation Prize Essay, 1980 (unpublished).
4. Pontecorvo, B. *Sov. Phys. JETP* 26, 984 (1968).
5. For a review, see: Bahcall, J.N., *Rev. Mod. Phys.* 50, 881 (1978).
6. Reines, F., Sobel, H.W., & Pasierb, E. (unpublished).
7. Feynman, R., & Vogel, P. (unpublished).
8. For a review, see Wachsmuth, H., CERN-EP/79-115C (1979). For an interpretation of these results in terms of neutrino oscillations, see De Rújula, A., Lusignoli, M., Maiani, L., Petkov, S.T., & Petronzio, R. *Nuclear Physics B* 168, 54 (1980).

A. De Rújula is at CERN, Geneva and S.L. Glashow is Professor of Physics at the Lyman Laboratory, Harvard University.

produced by the mechanisms that power the sun.⁵ These solar neutrinos have been detected on Earth when they transmute chlorine atoms to argon atoms. The observed flux of solar neutrinos appears to be too small by a factor of about three when compared with theoretical expectations. There are two popular explanations for this 'solar neutrino problem'. One stems from the observation that the chlorine experiment is sensitive to only a small minority of the solar neutrinos, the most energetic. These arise from a rare solar nuclear reaction, for which the calculation of the solar flux may be wrong. Accordingly, it is necessary to repeat the experiment with a detector sensitive to a broader range of neutrino energies. Many tons of such exotic and costly materials as gallium or indium are needed for this. But the phenomenon of neutrino oscillations could provide a more intriguing resolution of the solar neutrino problem. With maximal neutrino mixing, the solar neutrinos may have become an equal mix of the three neutrino species by the time they have reached the Earth. Since the method of detection of solar neutrinos is sensitive only to electron neutrinos, the apparent flux is thereby reduced to one-third. This explanation of the solar neutrino deficit requires neutrinos to have masses of at least 10^{-6} eV and to mix to a considerable extent.

Several recent ground-based experiments point in the direction (but do not amount to a compelling proof) of neutrino oscillations. Two of these tentative sets of results must be regarded as teasers for the more precise experiments that can and must be done.

Electron antineutrinos ($\bar{\nu}_e$) from a reactor were observed in a deuterium (d) detector eleven meters away.⁶ Two reactions induced by the antineutrinos were studied, $\bar{\nu}_e + d \rightarrow n + n + e^+$ and $\bar{\nu}_e + d \rightarrow n + p + \bar{\nu}$. The first is a conventional 'charged current' weak interaction to which only electron antineutrinos can contribute. Other antineutrinos can make only μ^+ or τ^+ , for which there is not enough energy. Thus, the first reaction would be inhibited by neutrino oscillations. The second reaction is a 'neutral current' weak interaction characteristic of the now accepted electroweak theory. It is unaffected by neutrino oscillations because the neutral currents involve all neutrino species equally. The ratio of the number of charged-current events to the number of neutral-current events is therefore sensitive to the existence of neutrino oscillations but relatively insensitive to the poorly-known flux of antineutrinos from the reactor.

Reines *et al.*⁶ measured this ratio and find that it is about half of what one would expect with no neutrino oscillations. This suggests considerable neutrino mixing and neutrino mass splittings of at least 0.5 eV. The statistics of the effect are not yet convincing, however, and the authors

cautiously conclude "... that further experimentation with $\bar{\nu}_e$ at reactors ... would be helpful in elucidating the effect here observed, as well as experiments at accelerators and with cosmic rays." The interpretation of this experiment has been criticized by Feynman and Vogel⁷, who have pointed out a possible inconsistency between the deuterium experiment and an earlier hydrogen experiment by the same group done under similar experimental conditions. When this inconsistency is taken into account, the evidence for neutrino oscillations becomes less than convincing. The controversy is alive, and a revised version of the Reines *et al.* manuscript is in circulation.

The second set of data comes from an accelerator experiment at CERN with a beam of neutrinos expected to consist of equal numbers of electron and muon neutrinos. These neutrinos are very energetic, so that the muon neutrinos may interact in the detector to make muons just as readily as the electron neutrinos may interact to make electrons. If there are no neutrino oscillations, one should therefore see equal numbers of electrons and muons in the detector. The same should be true if there are oscillations connecting electron and muon neutrinos, since as many ν_e become ν_μ as ν_μ become ν_e . The Big European Bubble Chamber, sitting in this neutrino beam, has so far observed about 20 neutrino-induced events⁸. Only half as many electrons as muons were seen, but the statistics are clearly limited. If the

observations are significant, however, a possible explanation is that electron neutrinos mix with tau neutrinos. This could explain the apparent systematic deficit of the electron neutrino component in the CERN experiment, in that of Reines *et al.* and in the solar neutrino experiment.

There are thus several suggestive experimental indications that neutrinos have mass and do oscillate. None of the experiments is beyond question. Yet, such results would fit quite well into current theoretical developments concerning the synthesis of the strong, weak and electromagnetic interactions. All of these 'grand unified theories' predict that baryon number is violated, so that all nuclear matter ultimately decays. Many of them also demand that neutrinos have mass, but say nothing of what these masses are. Particle theorists may not know quite what they are doing, but they do feel they are on the right track. Positive results from proton-decay experiments and from neutrino-oscillation experiments will tell us much about the nature of the underlying theory. For example, a model with neutrino masses involves seven more measurable parameters than a model with massless neutrinos. We need these hints. Moreover, proton-decay and neutrino-oscillation experiments do not depend upon the construction of ever larger accelerators. The high-energy frontier is certainly of fundamental importance, but there is still much to be learned at lesser energies. □

Esoteric DNA?

from Mervyn J. Turner and John S. Cordingley

ONE of the more esoteric areas within the field of mitochondrial DNA research has been the study of kinetoplast DNA (kDNA). Protozoa of the order Kinetoplastida, which includes the causative agents of the leishmaniasis and trypanosomiasis, characteristically contain a single large mitochondrion within which is found the kinetoplast, a disc-shaped mass of tightly packed DNA spatially associated with the basal body of the flagellum. Electron microscopy of spread kDNA shows a complex network of two kinds of interconnected DNA circles. The size of the smaller 'minicircle' component is species specific, varying between 700 and 2,300 base pairs, with each network containing about 10,000 minicircles. Each network also contains up to 50 copies of a larger 'maxicircle' component of up to 35 kilobase pairs. RNA transcripts of the maxicircle can be

readily detected and this is one of several pieces of evidence suggesting that maxicircle DNA is equivalent to the mitochondrial DNA of other eukaryotes.

There is no similar consensus on the function (if any) of the minicircles. A number of laboratories have suggested from hybridisation kinetics and restriction endonuclease digestion that the minicircles within each network are heterogeneous with respect to sequence. Using cloned minicircles from *Trypanosoma brucei*, Donelson *et al.* (*Plasmid* 2, 572; 1979) extended these data to show about 25% sequence homology between *T. brucei* minicircles. Chen and Donelson (*Proc. natn. Acad. Sci. U.S.A.* 77, 2445; 1980) now report the complete sequences of two cloned *T. brucei* minicircles, each of about 1,000 base pairs. The two molecules share about 12% homology within a nearly continuous stretch, and another 15% scattered randomly throughout the rest of the two sequences, thus confirming the prediction of about 25% homology.

Does this bring us any closer to an understanding of the function of the

Mervyn J. Turner and John S. Cordingley are in the MRC Biochemical Parasitology Unit, The Moltano Institute, Cambridge.

minicircle? Fouts and Wolstenholme (*Nucl. Acids Res.* **6**, 3785; 1979) found an RNA transcript of 240 bases from minicircles of *Crithidia acanthocephali*, but as yet no such transcript has been found in *T. brucei*. Intriguingly, there are potential coding transcripts of comparable size in each minicircle sequence, and in each case such potential transcripts begin within the 12% of continuous sequence homology. Circumstantial evidence suggests that this 12% homology is conserved within all minicircles of *T. brucei*, since it has been reported that the restriction endonuclease Taq I cleaves all *T. brucei* minicircles and a recognition sequence for this enzyme is found within the 12% homology (the only Taq I site in one of the kDNA sequences). There are no obvious origins of replication in the kDNA sequences, which are characterised by a distinct lack of palindromes and inverted repeats. Outside the 12% continuous homology, the sequences are remarkably A-T rich (77%), perhaps analogous to the high A-T content in non-coding

regions of other mitochondrial genomes.

These data leave open the question of minicircle function. Possible suggested functions for the kinetoplast, some more far-fetched than others, have included a structural role at the base of the flagellum; a means of segregating maxicircles equally at division; a repository of transposons; and an involvement in the control of antigenic variation or other gene expression.

Minicircles have many features in common with satellite and repetitive DNA in other systems. In addition to the presence of multiple copies, they possess a buoyant density distinct from the nuclear DNA and show internal repeats and sequence divergence. This feature is particularly obvious in *T. cruzi* minicircles, which appear to comprise four tandem repeats of a related sequence (Leon *et al. Biochem. Biophys. Acta* **607**, 221; 1980). Thus, having no demonstrated phenotypic effect, one may speculate that minicircles are an example of mitochondrial parasitic DNA! □

local differences in the frequencies of chromosome inversions.

Habitat selection is also involved in the chromosomal variation of the mosquito *Anopheles arabiensis*¹¹; in Nigeria mosquitoes possessing certain inversions are more likely to rest indoors (and hence to bite humans) than are others from the same population with a different chromosome constitution. However, not all inversion polymorphisms involve habitat selection. Atkinson and Miller¹² find no evidence that *D. subobscura* from particular habitats tend to return to them. This contrasts with *D. persimilis* but may be related to the evolutionary inflexibility of *subobscura*'s inversion polymorphism, which shows few geographic or temporal trends.

Nearly all animals are highly variable for the genes controlling soluble enzymes. Remarkably, there is now evidence that individuals with different genotypes for certain enzymes may select a habitat which increases their fitness. The freshwater isopod *Asellus aquaticus* lives on decaying leaves and is polymorphic at an amylase locus. Homozygotes for one of the amylase alleles select willow when given a choice between willow and beech in the laboratory to a much greater extent than do the other genotypes¹³. This is reflected in local differences in gene frequency in ponds containing both types of leaf. The alcohol dehydrogenase polymorphism of *D. melanogaster* is the most extensively studied of all enzyme polymorphisms. When larvae are placed at a junction between medium containing 17% ethanol and medium without ethanol, those homozygous for the Adh^F allele tend to move onto the alcohol medium; Adh^S homozygotes show no habitat preference¹⁴. This behaviour is related to the higher activity of Adh^F in metabolising ethanol and to the selective advantage conferred on flies bearing this allele when faced with an alcohol stress¹⁵. Clearly, a general tendency for animals of different enzyme genotype to select the ecological niche to which they are biochemically most suited might help to answer the vexed question of why there is so much molecular polymorphism in natural populations.

The theory of sympatric speciation — that a single species may split into two reproductively isolated populations

Can genes choose habitats?

from J.S. Jones

ALTHOUGH the history of evolutionary genetics is largely a tale of attempts to explain genetic variation there is no consensus of opinion as to the evolutionary forces involved. One possibility is that variation is promoted by the complexity of the environment; if one allele is favoured in one microhabitat and its alternatives in others, then natural selection within the population could maintain genetic variation.

There are a number of convincing instances of an association between the genetic diversity of a population and the ecological complexity of its habitat. In the genus *Drosophila*, for example, species living in the complex environment of a tropical forest have up to ten times as many chromosome inversions as those living in simpler habitats such as deserts or offshore islands¹. Again, studies of populations of *Drosophila willistoni* show there is a clear relationship between the extent of chromosome polymorphism and the floral diversity of the environment².

Since the first attempts to provide a theoretical basis for these effects were made by Levene³, many mathematical models of natural selection in heterogeneous environments⁴ have been put forward. In a recent paper, however, Maynard Smith and Hoekstra⁵ show that most of these models are limited in their usefulness; only under certain very specific conditions of the relative fitness of each allele in each ecological niche can they

maintain polymorphism. They therefore have little more value as a general explanation of genetic variation than do the older theories in which heterozygotes have an overall advantage over homozygotes⁶.

However, models of natural selection in complex environments become much more powerful if individuals of different genotypes are able to select the niche in which they are most fit⁷. Habitat choices linked to genetic differences have now been found for a variety of genetic polymorphisms.

The early experiments of Kettlewell in which melanic and non-melanic *Biston betularia* were released into a barrel painted with black and white stripes showed that the moths chose resting sites against which they were best camouflaged. This kind of experiment has now been extended to other species and to natural populations⁸. *Anolis* lizards in the Caribbean show a polymorphism related to camouflage in which the different morphs choose the appropriate habitat; unstriped individuals spend most of their time on the ground, while those with stripes prefer to live in the bushes in which they are best concealed⁹. Taylor and Powell¹⁰, working in the Sierra Nevada mountains of California in a locality which has a mosaic of woodland and open habitats, marked *D. persimilis* captured from the various habitats with fluorescent dusts. When they released them into the wild the flies showed a strong tendency to return to a habitat similar to the one in which they were first captured. This behaviour is associated with

1. Spieth, H.T. & Heed, W.B. *Annu. Rev. Ecol. Syst.* **3**, 269 (1972).
2. DaCunha, A.B. & Dobzhansky, Th. *Evolution* **8**, 119 (1954).
3. Levene, H. *Am. Nat.* **87**, 331 (1953).
4. Ewing, E.P. *Am. Nat.* **114**, 197 (1979).
5. Maynard Smith, J. & Hoekstra, R. *Genet. Res.* **35**, 45 (1980).
6. Lewontin, R.C., Ginzburg, L.R. & Tuljapourkar, S.D. *Genetics* **88**, 149 (1978).
7. Taylor, C.E. *Genetics* **83**, 887 (1976).
8. Boardman, M., Askew, R.R. & Cook, L.M. *J. Zool.* **172**, 343 (1974).
9. Schoener, T.W. & Schoener, A. *Evolution* **30**, 650 (1976).
10. Taylor, C.E. & Powell, J.R. *Oecologia* **37**, 69 (1978).
11. Coluzzi, M., Sabatini, A., Petrarca, V. & DiDeco, M.A. *Nature* **266**, 832 (1977).
12. Atkinson, W.D. & Miller, J.A. *Heredity* **44**, 193 (1980).
13. Christensen, B. *Heredity* **87**, 21 (1977).
14. Cavener, D., *Behav. Genetics* **9**, 359 (1979).
15. Clarke, B. *Genetics* **79**, 101 (1975).
16. Bush, G.L. *A. Rev. Ecol. Syst.* **6**, 339 (1975).

J.S. Jones is in the Department of Genetics and Biometry, University College, London.

without complete geographic isolation — has been put forward and discredited so often that Mayr has compared it to the Hydra which grew nine heads whenever one was struck off. Habitat selection might, nevertheless, lead to the genetic division of a population if the different genotypes mate within the micro-habitats which they have chosen. The most convincing examples of this are found in insects which parasitise plants and which have strong host-plant fidelity¹⁶. In Wisconsin the introduced hawthorn fly *Rhagoletis pomonella* showed a sudden — and presumably genetic — habitat extension to apples in 1864 and to cherries in 1960. These populations are now partially reproductively isolated from each

other so that at least the early stages of speciation have been completed. How widespread this mechanism is remains to be seen.

Most new ideas in evolutionary biology can be traced directly to the writings of Charles Darwin and habitat selection is no exception. In the *Origin of Species* he wrote: "The more diversified the descendants from any one species become . . . , by so much will they be better able to seize on many and widely diversified places in the polity of nature and so be enabled to increase in numbers." With evidence that choice behaviour may play a part even in molecular evolution, Darwin's suggestion is at last receiving the attention which it deserves. □

Switching on the muscle genes

from Harriet Harris

THE differentiation of muscle fibres from precursor myoblasts is accompanied by particularly striking changes in cellular morphology and function. Myoblasts are developmentally committed cells, but mononucleate and lacking the biochemical and structural attributes of muscle. Myofibre differentiation proceeds by fusion of myoblasts into long, multinucleate myotubes, and this is accompanied by pronounced changes in the pattern of protein synthesis, most notably a several hundred-fold increase (Devlin & Emerson *Cell* 13, 599; 1978) in the synthesis of actin, myosin and other myofilament proteins which assemble into sarcomeres. After some days, cultured myotubes are contractile and resemble terminally differentiated muscle fibres.

A central question concerning the mechanism of differentiation in this, or any, system, is how, and at what level are changes in gene expression regulated? Are the changes in protein synthesis a direct reflection of changes in the transcription of genomic DNA into RNA, or are there additional controls at the post-transcriptional level, that is, the processing of nuclear RNA into cytoplasmic messenger and its translation into protein? The answer may not be simple: multiple modes of regulation are indicated for muscle by a quantitative study by Devlin and Emerson (*op. cit.*) on the pattern of post-fusion protein synthesis. Interestingly, expression of the major muscle proteins myosin (heavy and light chains), troponin and tropomyosin appears to be coordinately

regulated, but a number of other muscle-specific, but mainly unidentified polypeptides exhibit different patterns of expression. Actin and myosin synthesis have, however, attracted the most attention and in the past, both transcriptional and post-transcriptional control mechanisms have been invoked to account for the immense proliferation of muscle myosin on cell fusion. More recent biochemical work (Strohman *et al.* *Cell* 10, 265; 1977) and *in situ* hybridization experiments (John *et al.* *Cell* 12, 501; 1977) on myosin mRNA show that, although some myosin message is present in pre-fusion myoblasts, the amount is greatly amplified in myotubes. This argues against the existence of a pool of stored message, but is consistent with either an increase in transcription on fusion, or an increase in the stability of message, as proposed by Buckingham *et al.* (*Proc. natn. Acad. Sci. U.S.A.* 71, 1466; 1974).

An account of a much broader attack on the whole area of gene expression in muscle differentiation is given in two recent papers from Gros' laboratory at the Institut Pasteur (Affara *et al.* *J. mol. Biol.* 141, 441; 1980; Affara *et al.* *J. mol. Biol.* 141, 459; 1980). These describe experiments to monitor changes occurring during differentiation in the widest possible range of gene products, at three levels: in the nuclear RNA (nRNA), that is, at the level of transcription, in the cytoplasmic mRNA population and in the pattern of polypeptides synthesized. Cultured cells representing three different stages in mammalian myogenesis were employed in this study: an undifferentiated mouse embryonal carcinoma cell line, a mouse teratocarcinoma-derived myoblast cell line and myotubes formed *in vitro*. To

compare the complexity of transcribed sequences at each stage of differentiation, Affara *et al.* hybridized excess total nRNA from each type of cell (undifferentiated, myoblast and myotube) to single-copy mouse DNA. The committed cells (myoblast and myotube) showed a 30% increase in the extent of hybridization over the carcinoma cells, implying that new DNA sequences are transcribed at this stage. The cytoplasmic mRNA (polysomal, poly A-containing mRNA) populations at each stage of differentiation were compared by hybridization experiments using complementary DNA (cDNA) probes for the mRNA from each cell type. It was thus shown that, although most mRNA sequences found in undifferentiated cells are still present in myoblasts and myotubes, at each stage in differentiation a new set of mRNAs enters the polysomes. Is the appearance of these messages under transcriptional control? cDNA probes enriched for myotube-specific sequences, or those common only to myoblasts and myotubes, were used to search for these sequences in nRNA. The results suggest that RNA found specifically in the polysomes of differentiated cells is absent from the nRNA of undifferentiated cells, implying that, in these cells, the corresponding genes are not transcribed. This conclusion was supported by testing the susceptibility of common and myotube mRNA-specific sequences in chromatin to nuclease attack; unexpressed genes are expected to show greater resistance than those transcribed (Weintraub & Groudine *Science* 193, 848; 1976). Such an increased resistance was indeed found for myotube-specific genes in carcinoma cell chromatin, but not in myotube chromatin. It thus appears that the presence of new messages in the polysomes of myoblasts and myotubes is regulated at the level of transcription. However, because some RNA sequences common to myoblasts and myotubes, and present in similar amounts in the nuclei of both show increased abundance in myotube cytoplasm, the authors suggest that post-transcriptional factors may modulate the quantities of some polysomal mRNAs.

The next question addressed by Affara *et al.* is whether the new group of mRNAs which appear in the cytoplasm of myotubes are indeed translated, and whether they code for identifiable, muscle-specific proteins. The set of myotube-specific messages was purified by a hybridization technique and translated *in vitro*. Many of the polypeptides produced correspond to new proteins synthesized by myotubes *in vivo*. These included, most rewardingly, myosin light chains, muscle actin and troponin as major components, thus confirming that the expression of these proteins is regulated by the appearance of specific, new messages at cell fusion.

There are still some flaws in the experimental system of Affara and colleagues. Limitations in analytical

technique (the cDNA probes are not pure; certain messages may not be efficiently translated *in vitro*) make it difficult to trace the expression of any particular gene from transcription to translation with any confidence. An increase in experimental resolution will be required to distinguish the relative contributions of transcription and message stability to the rate of synthesis of specific proteins. A further criticism concerns the validity of the teratocarcinoma cell line as a model for normal muscle development. Although

most of the identified muscle-specific polypeptides produced by the myotubes were characteristic of fast skeletal muscle, the authors also noted production of embryonic myosin light chain and slow muscle troponin C, which implies that there are aberrations in the regulatory mechanisms of these cells. However, it is to be hoped that the experimental approach outlined in these papers will lay the foundations for a more detailed understanding of the mechanisms of myogenesis. □

Multiphoton ionization interference effects

from Peter Knight

DIFFERENCES between results and theory in multiphoton ionization experiments seem to have been resolved by a recent experiment of the Saclay group (Morellec *et al.* *Phys. Rev. Lett.* **44**, 1394; 1980).

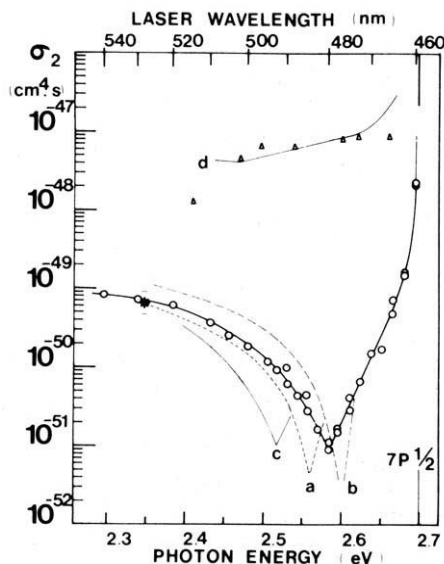
The ionization of atoms by the simultaneous absorption of several photons from an intense laser field has been the subject of intensive investigation recently, spurred on in part by the potential application to isotope separation. However, experiments and theory have been disagreeing by up to four orders of magnitude (*Nature* **270**, 561; 1977) even for the simplest example of non-resonant two-photon ionization of a simple alkali such as caesium. Calculations employing second-order perturbation theory all agree in predicting a deep interference minimum in the two-photon ionization cross-section of caesium at photon energies between 2.3 and 2.6 eV arising from cancellations of 6p and 7p intermediate state contributions. Experiments of Van der Wiel and colleagues (*J. Phys.* **B8**, 1617; 1975) saw no such minimum, and their observed cross-sections were four orders of magnitude greater than expected from theory. Several attempts have since been made to explain this anomaly and the experiment of Morellec *et al.* seem to have resolved the problem by using a very much more intense laser than that of Van der Wiel.

Most workers in the field have suspected that an explanation of the discrepancy between theory and experiment was connected with the presence of a small number of alkali dimer molecules in the atomic beam target. To find the interference minimum is difficult experimentally: the atomic cross-section is (obviously) very small. On the other hand caesium molecules can ionize and produce a detector signal very easily in the

wavelength range where the predicted atomic minimum should occur, since they possess many intermediate levels at the appropriate energies and can be reasonably ionized. The molecules are in fact so easy to ionize that they can dominate the ionization rate even if their concentration is very low (a few percent) of the atomic concentration. Van der Wiel and his coworker did attempt to isolate and eliminate molecular contributions yet still found their anomalously high cross-sections survived.

The Saclay group have now attacked the dimer problem in two ways. First, they work at much higher intensities than Van der Wiel using a three-stage oscillator-

Two-photon ionization cross-section of Cs atoms as a function of photon energy and laser wavelength (from Morellec *et al.* *Phys. Rev. Lett.* **44**, 1396; 1980). Triangles: experimental points of Van der Wiel normalized to new photoionization cross-sections; circles: experimental points of Morellec *et al.*; curves *a*, *b* and *c* are perturbative calculations noted in the text; curve *d* is a theoretical result incorporating laser bandwidths in a simple model of



amplifier dye laser pumped by the second-harmonic of a ruby laser. Their experiments could then be carried out using laser intensities of 10^{10} W/cm² compared with 10^5 W/cm² of Van der Wiel. Under such circumstances the resonant molecular ionization signal is saturated and is proportional to the intensity whereas the unsaturated atomic signal depends quadratically on the laser intensity and is easily distinguished from the molecular background. Secondly, the Saclay group performed their experiment with and without a caesium atomic beam superheater which efficiently dissociates dimers, and clearly see a reduction in the molecular component by more than an order of magnitude when the superheater is used. Using these two improvements together results in the atomic signal dominating the unwanted molecular signal in the 10-20 GW/cm² laser intensity range.

The Saclay results for the ionization count are converted into a value for the two-photon generalized cross-section σ_2 taking into account well-known multimode coherence effects. Their results for σ_2 are shown in the figure and compared with earlier results of Van der Wiel and various theoretical values. It is clear that the expected ionization minimum has been observed, with a reduction by two orders of magnitude in σ_2 in going from photon energies of 2.3 to 2.58 eV, and then a huge increase as the photon energy approaches the $7p_{1/2}$ intermediate-state resonance. Curves *a*, *b* and *c* are all derived from second-order perturbation theory using different atomic calculational methods: curve *a* is due to Bebb (*Phys. Rev.* **149**, 26; 1966), curve *b* to Teague *et al.* (*Phys. Rev.* **A14**, 1057; 1976) and *c* to Rachman *et al.* (*Phys. Lett.* **68A**, 433; 1978). These second-order perturbative theories are in quite good agreement with each other and with the Saclay data, and disagree with the data of Van der Wiel *et al.* Curve *d* in the figure, due to Armstrong and Eberly (*J. Phys.* **B12**, L291; 1979), although agreeing with Van der Wiel's data, disagrees with the new results and the second-order theoretical results. Curve *d* is obtained by incorporating laser fluctuations into the ionization dynamics in a first approximation using a phase-diffusion model of the laser bandwidth. This clearly exaggerates the laser bandwidth in the wings of the laser line and over-emphasises fluctuations in a way anticipated by Armstrong and Eberly who are currently working on improved laser-bandwidth models. The conclusion is that for the system studied by the Saclay group, second-order perturbation theory is adequate, that the interference minimum is observed once precautions are taken to eliminate molecular background, and that the earlier experiments at lower intensity were probably limited by this molecular background. It looks as if the non-resonant multiphoton ionization of simple atoms is now fairly well understood. □

Peter Knight is in the Optics Section, Blackett Laboratory, Imperial College, London.

The source of cells for regeneration

from J.M.W. Slack

A solution to the long standing problem of the source of cells for regeneration has been brought nearer by recent studies on planaria and the amphibian limb. The experiments use two techniques: X radiation doses which can inhibit cell division and regeneration without killing the tissues, and markers which allow the fate of cell populations to be followed.

Both planaria and the limb undergo 'epimorphic' regeneration in which a blastema consisting of cells which are visibly undifferentiated and indistinguishable from one another is formed at the wound site. The blastema cells first proliferate and then differentiate to reform the missing body part containing several different cell types.

The key questions are:— (1) Is the blastema formed (a) from cells drawn from all parts of the body, or (b) from tissues local to the wound? (2) Is the regenerate formed (a) by the *de novo* differentiation of reserve cells, or (b) by dedifferentiation and redifferentiation of functional tissue cells? (3) Is the blastema (a) a mixture of cells pre-committed to differentiate into particular cell types, or (b) composed of similar pluripotent cells able to become any of the cell types in the regenerate?

Although the questions are logically independent the discussion usually becomes polarised into an argument of 'neoblasts *versus* metaplasia'. 'Neoblast' is a term often used to indicate small basophilic cells present in planaria and other organisms, or sometimes in a special sense to indicate mobile pluripotent cells reserved for regeneration. 'Metaplasia' is the dedifferentiation of functional cells followed by redifferentiation of the cells or their progeny into a different histological type — for example transformation from myotube to chondrocyte. The 'neoblast *versus* metaplasia' argument thus corresponds to options 1a, 2a and 3b *versus* options 1b, 2b and 3b. Although the controversy has become fixed in this dichotomy, it must be remembered that other combinations of answers to the three questions are also possible.

J.M.W. Slack is in the Imperial Cancer Research Fund Laboratories, Mill Hill, London.

Planaria can regenerate from either an anterior or a posterior cut surface and contain scattered small basophilic cells which have been called neoblasts by many authors. Early experiments¹ showed that regeneration could be prevented by X-irradiation, but if the worm contained an unirradiated portion (either a shielded region or an unirradiated graft) then it could regenerate fully, even when the healthy portion was some distance from the level of amputation. The time taken for regeneration to begin increases with this distance and correlates with migration of the neoblasts through the intervening tissue.

This showed that a blastema can be formed from migratory cells and is capable of forming all the tissues of the regenerate. However, it did not exclude the possibility that the migratory cells were formed by dedifferentiation, or that in unirradiated worms differentiated tissues local to the wound can also contribute to the blastema. Indeed, histological and electron microscopic studies suggest that local dedifferentiation does occur².

Now studies have been published³⁻⁶ on the regeneration of a strain of planaria which are naturally occurring mosaics: the somatic cells are triploid, the cells of the male germ line are diploid and the cells of the female germ line are hexaploid. Regeneration from the gonadless regions gives rise to worms whose somatic tissues consist wholly of triploid cells, whilst regeneration from a cut surface through the gonadal region gives rise to blastemas and regenerates containing some diploid and/or hexaploid cells. Of particular significance is that myotubes with diploid nuclei were found in the muscle of the regenerated pharynx⁶. If germ cells and their precursors are accepted as *bona fide* differentiated cells then metaplasia has clearly occurred by the sequence: germ cell → blastema cell → muscle cell. However, it could be argued that the germ cells are a special case because in other circumstances they may exhibit pluripotency without passing through a sexual cycle, for example in mammalian teratocarcinoma. This objection might be overcome by building up stocks of worms

whose somatic tissues are of an unusual ploidy so that metaplasia between different somatic tissues can be looked for.

Many species of urodele amphibia will regenerate limbs after amputation. The limb contains no small basophilic cells and the histology of the blastema suggests formation by dedifferentiation of muscle, cartilage and connective tissue. Amputation through a small X-irradiated region does not result in regeneration even when the remainder of the limb and the rest of the animal is unirradiated⁷ so, in contrast to the planaria, it seems that a blastema cannot be formed without the participation of tissues local to the wound.

Recently studies on metaplasia have been carried out using axolotl tissues labelled by tritiated thymidine and triploidy⁸⁻¹⁰ or pigmentation and allogenic differences leading to graft rejection^{11,12}. Grafts of a particular marked tissue into an X-irradiated limb followed by amputation of the limb through the graft leads to the formation of a regenerate wholly composed of graft-derived cells, confirmed by the retention of the marker in the various cell types. The irradiated background tissue is necessary because although it does not contribute any cells to the regenerate it does contribute pattern information without which few structures are formed from a small graft¹³.

The results are summarised in the Table below. It is clear that there is extensive metaplasia between the different internal tissues, which are all derived from the mesoderm of the embryo. However, epidermis never contributes to the internal tissues and vice versa, indicating the existence of a lineage restriction between tissues of different embryological origin.

So although there may be some restrictions on the possible inter-conversions, there is now good evidence from both planaria and the amphibian limb that some functional cell types can dedifferentiate and that their progeny can appear in the regenerate as different histological types. These results should be taken seriously by those who advocate theories of cellular differentiation based on irreversible events such as the chemical modification or rearrangement of DNA. □

Cell types found in regenerates formed from grafts of particular tissues.

cell types in regenerate:	Grafted Tissue					
	ectodermally derived		mesodermally derived			
	epidermis	dermis	muscle	cartilage	nerve	(sheath)
epidermal	+	—	—	—	—	—
myotube	—	+	+	?	+	+
chondrocyte	—	+	+	+	+	+

+ marked cells are present, — marked cells absent, ? conflict of results between laboratories.

1. Wolff, E. in *Regeneration* (ed. Rudnick, D.) 53-84 (Ronald Press Co., New York, 1962).
2. Hay, E.D. in *The Stability of the Differentiated State* (Springer-Verlag, Berlin/Heidelberg, 1978).
3. Gregnigni, V. & Puccinelli, I. *J. Exp. Zool.* **199**, 57 (1977).
4. Gregnigni, V., Miceli, C. & Puccinelli, I. *J. Embryol. Exp. Morph.* **55**, 53 (1980).
5. Gregnigni, V., Miceli, C. & Picano, E. *J. Embryol. Exp. Morph.* **55**, 65 (1980).
6. Gregnigni, V. & Miceli, C. *Wilhelm Roux's Archives* **188**, 107 (1980).
7. Butler, E.G. & O'Brien, J.P. *Anat. Rec.* **84**, 407 (1942).
8. Steen, T.P. *J. Exp. Zool.* **167**, 49 (1968).
9. Namewirth, M. *Dev. Biol.* **41**, 42 (1974).
10. Dunis, D.A. & Namewirth, M. *Dev. Biol.* **36**, 97 (1977).
11. Wallace, H. *J. Embryol. Exp. Morph.* **28**, 419 (1972).
12. Maden, M. & Wallace H. *Acta Embryol. Exp.* **2**, 77 (1975).
13. Maden, M. *J. Embryol. Exp. Morph.* **50**, 325 (1979).

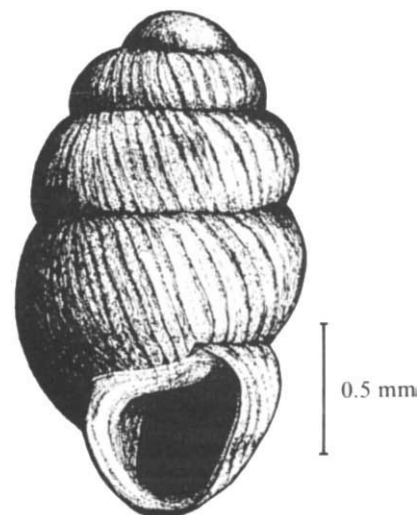
A glacial relict mollusc

from B. Coles* and B. Colville

THE mollusc *Vertigo genesii* (Gredler), thought to be extinct in Britain, has been found living in the Upper Teesdale National Reserve. *V. genesii* has a scattered distribution in Boreal and Arctic Europe where it lives in base-rich flushes of montane character^{3,4} and is common in British deposits of late-glacial age (12,500 — 10,000 BP)^{5,6}.

The Upper Teesdale National Nature Reserve, UK, is well known for its exceptional flora, including a relict

arctic-alpine element.¹ A similar faunal element has not been recognised although the Reserve contains invertebrates of rare occurrence in Britain.² The specimens were found, alive, in tufts of *Carex demissa* Hornem. (agg.) in the open parts of base-rich flushes on Widdybank Fell (450 m) in May 1979; their presence was confirmed in July of the same year. Mollusca associated with *V. genesii* were:— *Lymnaea truncatula* L., *Succinea pfeifferi* (Rossmässler) and *Nesovitrea hammonis* (Ström). Plants present were:— *Armeria maritima* Willd., *Plantago maritima* L., *Juncus triglumis* L., *Kobresia simpliciuscula* (Wahlen.), *Carex demissa* Hornem. (agg.), *C. panicea* L. and *Equisetum variegatum* Web. & Mohr. *Minuartia stricta* Hiern (which is probably the most notable plant of the Teesdale flushes, Teesdale being its only station outside the arctic circle) was not found at this site but the flush is otherwise typical of areas in Upper Teesdale where an arctic-alpine flora occurs. The occurrence of *V. genesii*



Vertigo genesii

in Upper Teesdale and its association with the relict arctic-alpine flora agrees well with the European distribution of the species. It is noteworthy that the associated Mollusca in Upper Teesdale were also common associates of *V. genesii* in the late-glacial. □

1. For a summary of the natural history of Upper Teesdale see:— Clapham, A.R. Ed., Upper Teesdale, Collins, London (1978).
2. See Coulson, J.N. in chapter 8 of ref. 1.
3. Ehrmann, P. Die Tierwelt Mitteleuropas II, Weichtiere, Mollusca, Quelle & Meyer, Leipzig (1933).
4. Zilch, A. and Jäckel, S.G.H. Ergänzung Tierwelt Mitteleuropas Weichtiere, Mollusca, Quelle & Meyer, Leipzig (1962).
5. Kerney, M.P. Late-glacial deposits on the chalk of South-East England, Phil. Trans. Ser. B 246, 203-254 (1963).
6. Kerney M.P., Brown, E.H. and Chandler, T.J. The Late-glacial and Post-glacial history of the chalk escarpment near Brook, Kent, Phil. Trans. Ser. B 248, 135-204 (1964).

We would like to thank the Nature Conservancy Council and in particular J. P. Doody, for permission to do work in Upper Teesdale National Nature Reserve.

*Author's address: 27 Dove St., York.

Sign, symbol and syntax in the language of apes

from Brendan O. McGonigle

TERRACE'S¹ recent criticism of evidence suggesting that apes can master "the conversational, semantic, or syntactic organization of language" has intensified a controversy spanning a century of speculation by primatologists. In the 19th century, several determined attempts were made to teach apes to talk but they were largely unsuccessful, only a few words being mastered by some animals. Later, Yerkes² identified part of the problem as the medium of communication. Apes would not mimic what they heard; instead their powers of imitation were heavily dependent on vision. Impressed by their intelligence, especially as a result of Köhler's now classic studies of insightful behaviour by chimpanzees³, Yerkes suggested that apes "might be taught to use their fingers, somewhat as does the deaf and dumb person, and thus helped to acquire a simple, non-vocal 'sign language'".

More than 40 years later this suggestion was put to the test⁴ when a young chimpanzee, Washoe, was trained by the Gardners on a form of manual-visual communication claimed to be a version of American sign language (ASL). By the third year of the project, Washoe had acquired a stable repertoire of some 85 signs of which a number were used in two sign combinations in a manner compatible

with that reported for children⁵. In addition, Washoe invented a number of signs of her own. A striking case was *bib* whose ASL equivalent was not known to her trainers when it was first introduced (appropriately) by the chimp. In later tests the Gardners monitored Washoe's capacity to answer 'wh' questions (who, what, where, whose) the production of which constitutes the first evidence that children have grammatical competence⁵. With various simplification procedures for analysing their data (now contentious) the Gardners found that the question frames exerted a very high control over the replies as indexed by the grammatical appropriateness of the answers. In carrying out the tests the Gardners relied on productive (making the chimp formulate a reply in terms not given in the question) rather than comprehension tests which "effectively eliminates the problem of clever Hans cueing".

Despite its apparent scientific rigour, the work has had its critics. In a recent attack, Seidenberg and Petitto⁶ claim that Washoe's signing does not resemble signing in ASL in any important way. Instead, they allege that a large proportion of ape signs can be interpreted without any special knowledge of apes or ASL and form part of a repertoire of unlearned gestures. These authors also challenge the Gardners' method of simplifying or reducing the replies of Washoe to 'wh' questions before reporting them. Amongst those deleted

were signs which were repetitious of those in a question and signs which were repeated during a reply. The effect is to trim a response such as *you me you out me to you me* thus giving the utterance a more humanlike form — a possibly serious distortion in their view. By contrast, children in the early stages of language acquisition "typically produce utterances which are reduced relative to the corresponding adult forms, leading to the characterization of their speech as telegraphic".

Terrace and his associates have also found evidence for redundancy in the signing of their chimpanzee Nim, trained like Washoe in what is in fact Pidgin Sign English (PSE). Some utterances contained as many as 16 signs issued in "a striking, redundant, recursive string". In analysing the relationship between 2, 3, and 4 sign combinations, moreover, Terrace has found a major difference between Nim and children. When making longer utterances the child deletes repeated elements and expands on the information given in the shorter utterance. By contrast Nim's additions to the two sign utterances tended to be redundant, for example, "play me", "play me Nim" or repetitious, for example, "eat Nim eat". Terrace suggests that the purpose of three sign combinations by Nim is to add emphasis rather than convey new information.

The really significant findings in the report by Terrace *et al.* which

Brendan O. McGonigle is in the Department of Psychology, University of Edinburgh, Scotland.

indicate that Nim has little language ability are the following. (i) An increase in vocabulary to 125 signs during the course of the training programme did not result in a related expansion in the mean length of utterance (MLU). This remained fairly stable over the course of the programme. (ii) There was little originality in Nim's utterances. A high percentage (87%) were classifiable as 'adjacent' (i.e. following the teacher's utterance without a definite pause) and a large percentage of these (~40%) were either imitations or 'reductions' (containing some of the utterances of the teacher and nothing else). (iii) Nim showed little evidence of the 'turn-taking' strategy which characterizes human conversation. Instead a large majority of his signs overlapped with those of his teachers — most of them interruptions which could not be construed as attempts to 'take the floor'.

To check that Nim was not an atypical case, Terrace *et al.* have analysed two films of Washoe which also include signing behaviour by Ally (Nim's full brother) and Koko, a gorilla⁷. Their analysis shows they claim that "Washoe's utterances were adjacent and imitative of her teachers' utterances . . . ninety-two percent of Ally's, and all of Koko's, signs were signed by the teacher immediately before Ally and Koko signed".

The Gardners hotly deny that Terrace's analysis is accurate, claiming instead that single frame and slow-play analyses misrepresent and distort the record. They also draw attention to the fact that Washoe frequently signed to herself when she was unaware of an observer, or when looking at pictures in magazines⁴. As things stand it will be some time before the technical and the statistical issues raised by Terrace's analysis can be resolved.

One conclusion which does emerge, however, is that where there is suspicion that some of an ape's signs are merely imitative of an earlier use in the same test by the teacher, the use of syntactic criteria including word order may reduce the danger of 'over-attribution' by the experimenter. To use these criteria it would be necessary to establish that the constituents of any utterance had particular meanings when presented in isolation, that signs in different linear combinations had different meanings, and that each order structure is not specific to a unique combination of signs. So far none of the chimpanzee language projects appear to have met all of these criteria — at least within the same study. Premack⁸, who trained a chimp, Sarah, using an artificial language based on plastic symbols claims that "chimps can be taught word order but only with explicit training . . .". He found no evidence that Sarah could produce structural innovations, although he claims she converted the trainer's (production) order of symbols and "produced sentences by systematic rearrangement . . .". On this latter point,

however, he offers no evidence. As for word understanding, the use of too few alternatives in the tests employed frequently leads to ambiguity as to what the chimp has understood. (For instance, the preposition *on* was taught in isolation such that word order was always a simple clue to the subject-object relationship specified. It would not have been had this preposition been contrasted with *under*).

Stronger evidence for grammatical competence comes from Project Lana⁹ where chimps have been trained to use lexigrams from a language called Yerkish. A computer monitors the productions of the subject — achieved by depressing keys — which must conform to the rules of a correlational grammar before any requests can be honoured, e.g. *Please machine give chow period*. The words *please* and *period* are fixed features of the sentence frame to enable the computer to determine the beginning and the end of an utterance. Examination of the performance of the subject Lana over one month of training revealed that she composed on her keyboard "76 strings of 6 lexigrams that were grammatical sentences and that did not figure in any training programme, whereas during the same period she produced only 71, 6-lexigram strings that were ungrammatical"¹⁰. Clearly there is evidence here of innovation. However, the analysis did not include consideration of the communicative or the contextual appropriateness of these grammatical strings and it seems likely that not all of the constituent lexigrams in the strings carried an independent meaning for Lana. For example, *please* is an invariant feature of a sentence frame necessary to get the computer to accept a message. It is improbable, as Seidenberg and Petitto⁶ point out, that the appropriate depression of this key suggests that the ape has a sense of politesse. More recent work on this project has shown, furthermore, that apes begin to use individual lexigrams as referents for objects and actions after they are able to produce ordered sequences of them to produce food. The apparent

ordering of signs must be rigorously analysed to determine the extent to which the sequences are controlled simply by lexical order rules independent of their (constituent) meaning.

Issues of syntax apart, it seems pertinent to ask what apes tutored in these various 'languages' have revealed about their mentation which could not be inferred from other aspects of their behaviour. The results thus far are ambivalent. On the one hand the content of ape language merely reaffirms the narrow incentive range within which the chimp seems to operate in laboratory situations. His favourite words seem to be those which invite the listener to donate food or drink, or which command attention (although this pattern may change when more work is carried out with mature subjects). On the other hand, the restrictions on semantic role which Terrace describes in the case of Nim (indexed by a marked restriction on the range of agents and beneficiaries, e.g. NIM, me, you and the names of other animates) are symptomatic perhaps of an egocentric form of thought — a characteristic of early stages of child development according to Piaget — not easily revealed by non-verbal tests. The response of apes to attempts to teach them the spatio-temporal conventions for reference to objects and events outside the immediate context of utterance should provide further important clues to the cognitive resources available to these enigmatic primates. □

1. Terrace, H.S., Petitto, L.A., Sanders, D.J. & Bever, T.G. *Science* **206**, 891 (1979).
2. Yerkes, R.M. *Almost Human* (Jonathan Cape, New York, 1925).
3. Köhler, W. *The Mentality of Apes* (Liveright, New York, 1925).
4. Gardner, B.T. & Gardner, R.A. *J. exp. Psychol.* **104** 244 (1975).
5. Brown, R. *A First Language: the Early Stages* (George Allen and Unwin, London, 1973).
6. Seidenberg, M.S. & Petitto, L.A. *Cognition*, **7**, 177 (1979).
7. Patterson, F. *Brain and Language* **5**, 72 (1978).
8. Premack, D. *Intelligence in Ape and Man* (Lawrence Erlbaum Associates, Hillsdale, N.J., 1976).
9. Rumbaugh, D.M. *Language Learning by a Chimpanzee: The Lana Project* (Academic Press, New York, 1977).
10. Von Glasenberg, in *Language Learning by a Chimpanzee* (Ed. Rumbaugh, D.M.) (Academic Press, New York, 1977).

100 years ago

HUMAN HYBERNATION

In the *Transactions* of the Royal Society Dr. W. Oliver has recorded the history of an extraordinary sleeping person named Samuel Chilton of Tinsbury, near Bath, who on May 13 1694, being then 'of robust habit of body, not fat, but fleshy, and a dark brown hair,' happened, without any visible cause or evident sign, to fall into a very profound sleep, out of which no art used by those who were near him could rouse him until after a month's time; then he rose of himself, put on his clothes, and went about his business of husbandry as usual; slept, could eat and drink as before, but spoke not one word till about a month after. In 1696, on the 9th of April, this youth fell off to sleep again, and although a heroic apothecary, Mr. Gibbs, bled him, blistered him, and scarified him, he slept on for seventeen weeks, waking up on August 7, not knowing

he had slept above a night, and unable to be persuaded he had lain so long, until going out into the fields he found everybody busy getting in the harvest, and then remembered very well that when he fell asleep they were sowing of the barley and oats which he now saw ripe and ready to be cut down. For six weeks of this sleep he had fasted, but after he awoke he went to work in his ordinary way, and continued to work until August 17, 1697, when, after complaining of shivering and cold in his back, and vomiting once or twice, he fell into one of his long sleeps once more. So he lay sleeping until November 19, when he awoke, said he "felt very well, thank God," ate some bread and cheese, and dropping off still another time, slept on until the end of January, 1698, and "then waked perfectly well, not remembering anything that happened all this while."

REVIEW ARTICLE

Nucleosome assembly

Ronald A. Laskey & William C. Earnshaw

MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK

Histones and DNA can spontaneously associate to form the nucleosome subunits of eukaryotic chromatin, but two proteins which occur in the eukaryotic nucleus can facilitate nucleosome assembly and greatly extend the conditions which permit assembly to occur. Recently several laboratories have reported new experimental systems which help to analyse, reconstruct and exploit the cellular mechanisms of chromatin assembly.

IN eukaryotic nuclei DNA is packaged into regularly repeating nucleoprotein units called nucleosomes. The structure of the nucleosome has been reviewed extensively¹⁻⁴ and the major features of the prevailing model⁵ are summarized in Fig. 1. Two structural domains have been distinguished by nuclease digestion, a protected nucleosome core and an exposed linker DNA segment of variable length. The nucleosome core consists of 146 base pairs of DNA coiled around a central histone octamer containing two each of the histones H2A, H2B, H3 and H4. Histone H1 associates with a further 20 base pairs of DNA adjoining the core, to complete two full superhelical turns of 80 base pairs each around the histone octamer^{6,7}. In addition to this simple repeating structure chromatin possesses a complex range of local variations in the distribution of nonhistone proteins^{8,9}.

Studies of chromatin assembly have been directed both at determining the pathway of nucleosome assembly and also at analysing the functional significance of variations in chromatin structure. Chromatin assembly by dialysis against gradients of salt under nonphysiological conditions (starting at >1 M NaCl) has produced information about nucleosome structure, but attempts to attach biological significance to the assembled structures have been repeatedly confused by artefacts (reviewed in refs. 2-4).

Recently several new experimental systems for assembling nucleosome cores or nucleosomes at physiological ionic strength have been described. They fall into three classes: (1) self-assembly systems^{10,11} in which histones and DNA associate without assembly factors; (2) factor mediated assembly systems in which assembly is facilitated by purified assembly factors^{12,13}; and (3) unfractionated cell homogenates¹⁴⁻¹⁶ which attempt to mimic the cellular processes and which can be fractionated further to elucidate the cellular mechanisms of chromatin assembly. All three classes will be considered here, together with investigations of assembly of replicating chromatin in intact cells. At present only the unfractionated cell homogenates have been able to reconstitute the physiological spacing between adjacent nucleosomes.

Self-assembly of the nucleosome core

Histones and DNA contain the morphogenetic information required to assemble a nucleosome core. This has been clearly demonstrated by the reconstitution of nucleosome cores by salt gradient dialysis from approximately 1 M NaCl (refs 17-21).

Although the capacity for self assembly was readily demonstrated at unphysiological salt concentrations, it has only recently been observed at ionic strengths close to physiological. Ruiz-Carrillo *et al.*¹⁰ renatured core histones into H3/H4 tetramers and H2A/H2B dimers by prolonged dialysis and slowly pumped the histones into an excess of DNA at 0.2 M NaCl. The product was shown to be authentic nucleosome cores¹⁰ (Fig. 2). If histone H1 was included a precipitate formed. This could be redissolved by dialysing the precipitate to lower

salt concentrations (10-30 mM). In a second study by Stein *et al.*¹¹ the core histones and DNA were mixed more abruptly. A precipitate formed, but redissolved during 16 h at 37 °C, to reveal substantial formation of nucleosome cores. In both cases excess histones caused irreversible precipitation.

Although these two assembly procedures^{10,11} are limited in the conditions which yield soluble products rather than insoluble precipitates, they clearly establish the principle that nucleosome cores can self-assemble from histones and DNA in the absence of other factors.

Nuclear proteins which facilitate nucleosome core assembly *in vitro*

Two proteins from the eukaryotic nucleus allow nucleosome core assembly to proceed rapidly *in vitro* without precipitation. One appears to interact with DNA and the other with histones. They are respectively DNA topoisomerase I (nicking-closing enzyme)¹³ and nucleoplasmin, an acidic thermostable protein which was originally isolated from *Xenopus* eggs¹².

The assembly pathway mediated by DNA topoisomerase I involves making a complex between topoisomerase I and DNA and then adding histones¹³. Approximately one topoisomerase I molecule is required for each nucleosome assembled, corresponding to 20 times more of the enzyme than would relax the same weight of supercoiled DNA. Assembly is inhibited when the weight of histones exceeds the weight of DNA. It is not clear yet if topoisomerase I acts by decreasing the electrostatic attraction between DNA and histones (see below) or by performing some additional function such as providing a swivel. DNA topoisomerase I occurs at substantial concentrations in the chromatin of a wide range of eukaryotes. Its properties and distribution have been reviewed elsewhere²².

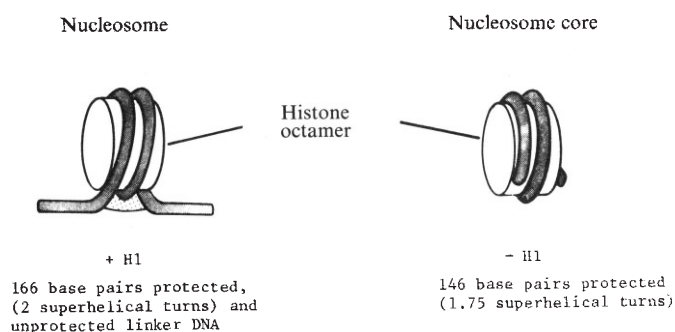


Fig. 1 A model for the structure of the nucleosome^{5,6} and nucleosome core⁵, in which DNA is wrapped around the surface of a flat protein cylinder consisting of two each of histones H2A, H2B, H3 and H4 (refs 1-4).

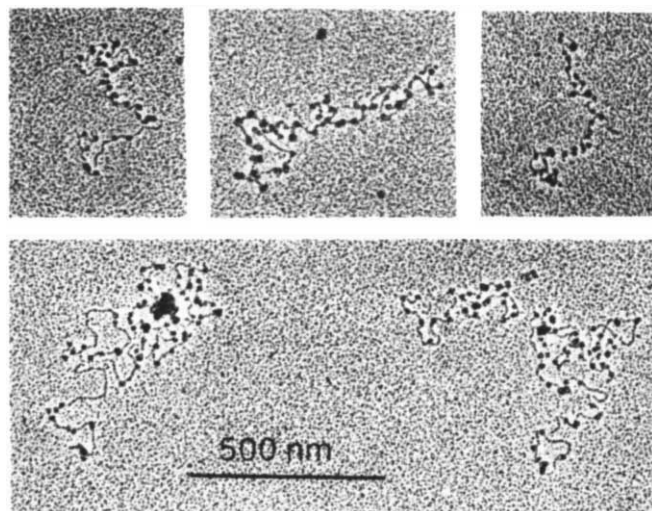


Fig. 2 Electron micrograph of nucleosome cores assembled by very slowly pumping core histones (without H1) into excess DNA at physiological ionic strength. Reproduced with permission from ref. 10.

Fractionation of a cell-free nucleosome assembly system¹⁴ from eggs of *Xenopus laevis* led to the identification and purification of an acidic, thermostable protein which facilitates nucleosome core assembly *in vitro* when histones and DNA are rapidly mixed¹². The protein is rich in potentially charged amino acids (notably 30% Asx + Glx)²³. It is phosphorylated^{24,25} and it has an isoelectric point of 5 (ref. 26). In solution it is an unusual multimer, a pentamer of the $M_r = 29,000$ –30,000 subunits, and this pentamer is active in nucleosome core assembly²⁷. The protein does not bind detectably to DNA or nucleohistone²⁷ but it does bind histones *in vitro*^{12,27}. Its activity is independent of DNA topoisomerase^{12,23,27}, and the properties of the protein, which itself has no associated topoisomerase activity, are strikingly different from known DNA topoisomerases²². Unlike DNA topoisomerase I it can mediate nucleosome core assembly in the presence of excess histones^{12,27}.

Recently this protein has been shown to be the most abundant protein in the *Xenopus* oocyte nucleus, occurring at about 5 mg ml^{-1} and representing 7.5–10% of the total nuclear protein^{24–26}. It is localized in the nucleoplasm^{24–26} and using a combination of immunofluorescence and immunoprecipitation, Krohne and Franke have shown that this, or a conserved, antigenetically related protein, occurs at substantial concentrations in the nucleoplasm of a wide range of vertebrate cell types²⁵ (Fig. 3). The probable reason why it had not been observed in nuclei previously is rapid leakage from nuclei in aqueous isolation buffers^{24–26}. Thus it was first discovered in total homogenates of cells in meiotic metaphase¹². With the agreement of Krohne and Franke we propose to call this protein 'nucleoplasmin' reflecting its unusual intracellular location.

Although it is not clear if nucleoplasmin is tightly bound to the histone pool of *Xenopus* oocytes^{24,25}, its ability to bind histones *in vitro* led to the suggestion that it assembles nucleosomes *in vitro* both by neutralizing electrostatic repulsion between histones thus promoting histone:histone interactions, and also by acting as a 'molecular chaperone' which competes against the strong electrostatic attraction between histones and DNA, thus minimizing nonspecific insoluble aggregates^{12,27}.

These interpretations are consistent with the observations of Stein *et al.*¹¹ that polyglutamic and polyaspartic acids can also greatly facilitate nucleosome core assembly when used in the conditions described for nucleoplasmin. In addition Stein *et al.*¹¹ showed that polyglutamic acid causes histones to associate to form stable octamers in the absence of DNA and at physiological ionic strength. These observations may partly explain the mechanism of nucleosome core assembly by nucleoplasmin *in vitro*.

Although nucleoplasmin assembles nucleosome cores *in vitro*, the possibility that *in vivo* it performs a wider role of mediating between the many highly charged components of the nucleus deserves investigation.

The methods for assembling nucleosome cores *in vitro* at physiological ionic strength are summarized in Table 1. From a comparison of the methods the following principles emerge. First the steric information for self-assembling a nucleosome core resides in histones and DNA themselves. Second nonspecific electrostatic interactions between histones and DNA or between histones and nucleosomes^{28,29} can result in insolubility of the product. This problem can be minimized by careful control of the experimental conditions, but it is accentuated by the presence of histone H1 or excess free histones. Third, the problem of solubility can be avoided by using 'assembly factors'. Fourth, these may act in different ways. Thus DNA topoisomerase interacts with DNA permitting subsequent ordered interaction with histones while nucleoplasmin and polyglutamic acid or polyaspartic acid appear to promote histone:histone interactions and do not interact directly with DNA. The simplest unifying hypothesis appears to be that the assembly factors described so far act *in vitro* by competing against or masking the strong electrostatic interaction between histones and DNA thus allowing self assembly to occur without aggregation and precipitation. Although it is possible that the assembly factors are recycled there is no evidence that they act truly catalytically or that they contribute steric specificity to the assembly pathways.

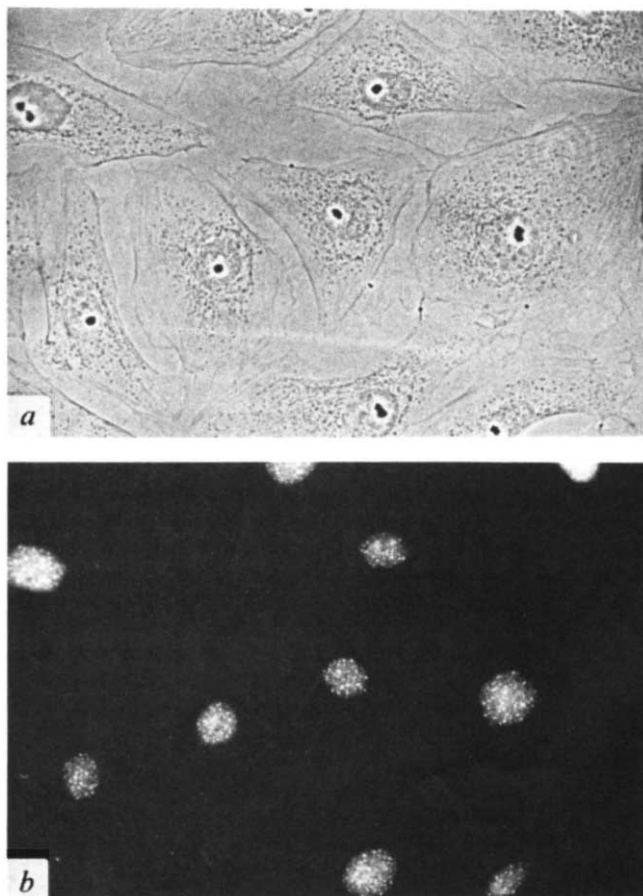


Fig. 3 Phase contrast (a) and immunofluorescence microscopy (b) of rat kangaroo cells incubated with antiserum raised against nucleoplasmin from oocyte nuclei of *Xenopus laevis*. The cells all show a strong and specific staining of their nuclei. Scale bars, 20 μm . Photograph supplied by G. Krohne and W. Franke (see refs 24 and 25).

Table 1 Methods for assembling nucleosome cores *in vitro* using purified reagents at physiological ionic strength

Method		Approximate weight ratios	Time to completion	Product characterization†	Reference to method
<i>Without assembly factors:</i>					
(1) slow pumped addition of histones to excess DNA		1 DNA ≤ 1 core histones	2–5 h*	A–E	10
(2) rapid mixing		1 DNA ≤ 1 core histones	16 h (precipitates and redissolves slowly)	A–C	11
<i>With assembly factors:</i>					
Factor	Normal location				
(1) DNA topoisomerase I	Chromatin (eukaryotes)	1 DNA ≤ 1 core histones 0.4 topoisomerase I	< 1 h	A–D	13
(2) Nucleoplasmin	Nucleoplasm (amphibians, birds, mammals, possibly elsewhere)	1 DNA 2 core histones 3–4 nucleoplasmin	< 1 h	{ A–C D–E	12 27
(3) Polyglutamic acid	Synthetic	1 DNA 2 core histones 3–4 polyglutamic acid	~ 5 min	A and F	11

* Precipitates when histone H 1 is present.

† Key to product characterisation: A, Electron microscopy; B, supercoiling; C, sedimentation; D, digestion by micrococcal nuclease; E, digestion by DNase I; F, protein cross-linking

Cellular mechanism of nucleosome assembly

There is no direct evidence yet that either DNA topoisomerase I or nucleoplasmin is involved in the cellular mechanism of nucleosome assembly. However, the following observations are consistent with their involvement. First, both proteins occur in the nucleus at concentrations sufficient to assemble nucleosome cores *in vitro*. Second, when naked DNA is introduced into *Xenopus* oocyte nuclei or into eggs which are in meiotic metaphase it is rapidly relaxed by high concentrations of topoisomerase before it is assembled into nucleosomes^{14,30}. Third, eggs of *Xenopus laevis* contain a large histone pool, sufficient for 20,000 diploid nuclei^{31,32} and of the assembly methods described above only nucleoplasmin (or polyglutamic and aspartic acid) allows nucleosome assembly and chromatin solubility in the presence of excess histones^{10,11}. Fourth, the endogenous histones in *Xenopus* egg homogenates exhibit anomalous charge properties, adsorbing to positively charged but not negatively charged resins at neutral pH suggesting that they are bound in a complex with a net negative charge. Purified nucleoplasmin mimics this effect when mixed with histones²⁷. Finally, the rate of nucleosome core assembly by purified nucleoplasmin is similar to the rate in unfractionated homogenates and is sufficient to account for the rapid rate of nucleosome assembly in the early *Xenopus* embryo²⁷.

Recent work of Nelson *et al.*¹⁶ suggests that *Drosophila* embryos may contain yet another nucleosome assembly factor. Although it has not yet been characterized, it seems to differ in some of its properties from either nucleoplasmin or DNA topoisomerase I. It will be interesting to see if it is related to a polypeptide, 'X', which Worcel *et al.*³³ observed associated with nascent nucleosomes in cultured *Drosophila* cells. In the same study³³ newly replicated DNA was observed to associate stably with H3 and H4 2–10 min before it associated stably with H2A and H2B and 10–20 min before it associated with H-1. A similar sequence of events has been deduced for mouse hepatoma cells³⁴ and replicating SV40 chromatin³⁵. Ruiz-Carrillo *et al.*¹⁰ have shown that a similar sequence of additions can be used to assemble nucleosomes *in vitro*. While this sequence of histone depositions may be widespread, it is not clear if it is essential *in vivo*. Thus the time course of histone depositions described by Worcel *et al.* is unlikely in early *Drosophila* embryos where the entire interphase of the cell cycle lasts only 3.4 min^{36,37} shorter than the lag between depositions of different histones observed by Worcel *et al.* From *in vitro* studies there is evidence that nucleosomes can be assembled by interaction of DNA with

either preformed histone octamers^{11,29} or dissociated dimers and tetramers¹⁰. It is possible that the multiple nucleosome assembly pathways *in vitro* reflect multiple alternative pathways in the cell, which would not be surprising if the problem of assembling nucleosome cores correctly involves simply limiting electrostatic interactions between histones and DNA to localized regions, nucleosomes, and preventing indiscriminate electrostatic cross-linking leading to precipitation.

An aspect of nucleosome assembly *in vivo* which we have not considered above is the role of transient post-translational histone modifications seen in newly synthesized histones^{38–40}. Since nucleosome cores can be assembled efficiently *in vitro* from unmodified histones (Table 1) it is tempting to speculate that the modifications might have an alternative role such as specifying the periodicity between adjacent nucleosomes in chromatin.

Assembly of regularly spaced chains of nucleosomes

Nucleosome cores are not distributed randomly along DNA, but are spaced at regular intervals of approximately 200 base pairs^{41,42}. The average length of the linker DNA varies between different cell types so that the nucleosome repeat length ranges from about 160 base pairs in fungi and rabbit cortical neurones to 241 in sea urchin sperm^{43–47}.

However none of the purified *in vitro* assembly systems listed in Table 1 has reproduced the native spacing. They either space nucleosomes randomly or at intervals which are shorter than those from the parent native chromatin^{10,13,27}. In contrast, as seen in Fig. 4, the unfractionated cell-free assembly system from *Xenopus* eggs spaces nucleosomes regularly at about 195 base pairs¹⁴ and the same has recently been shown for an unfractionated assembly system from *Drosophila* embryos¹⁶. The mechanism which specifies the length of the internucleosome linker is unexplained but some models can be firmly eliminated. First, the correct spacing does not arise by copying the spacing on a parental molecule. It can arise *de novo* on naked DNA^{14,16} (Fig. 4). Second, it does not require DNA replication or the discontinuous synthesis of Okazaki fragments since it can occur in the absence of DNA replication¹⁴. Third, it is unlikely to arise from a conserved periodic signal in DNA since spacing is observed on reconstituted prokaryotic templates (ref. 16 and W.C.E. unpublished observation) which are not naturally assembled into nucleosomes. On the other hand, it seems possible that it may be mediated by transient nucleosome sliding

since it occurs in circumstances where assembly appears to be non-cooperative¹⁴, suggesting that a mechanism must exist for correctly aligning randomly deposited nucleosomes. So far correct inter-nucleosome spacing has only been observed in assembly systems which use an endogenous stored histone pool, namely egg extracts from *Xenopus* and *Drosophila*^{14,16}. Woodland⁴⁵ has shown that the stored histones of *Xenopus* eggs have secondary modifications characteristic of newly-synthesized histones in somatic cells, which raises the possibility that the transient histone modifications may be the missing ingredient required to achieve correct spacing in purified assembly systems. This hypothesis can be tested, and with the emergence of evidence suggesting that nucleosomes have specific phase-relationships with some DNA sequences⁴⁹⁻⁵⁵ there is an obvious need to develop purified assembly systems which position nucleosomes precisely on DNA.

Relationship of nucleosome assembly to DNA replication

Nucleosomes persist in the region of replication forks and both daughter DNA duplexes are rapidly converted to nucleosomes⁵⁶⁻⁶⁰. By labelling nascent histones with dense amino acids and ³H-lysine, Leffak *et al.*⁶¹ have shown that parental histone octamers segregate conservatively to one of the daughter duplexes and that groups of parental nucleosomes segregate together. There is disagreement, however, over the site of assembly of new nucleosomes. Studies in which newly replicated chromatin is density-labelled by dense thymidine analogues, then fixed in formaldehyde and separated from bulk chromatin on the basis of its greater density have repeatedly suggested that new histones associate with unreplicated DNA⁶²⁻⁶⁴. In contrast, studies which have fractionated nascent chromatin from bulk chromatin by sedimentation avoiding the use of dense thymidine analogues or formaldehyde fixation have suggested that new histones are deposited at or near the replication fork^{33,35}. Additional support for the view that new nucleosomes form on newly replicated DNA comes from experiments in which protein synthesis is inhibited by cycloheximide⁶⁵. Here putative replication forks can be observed in

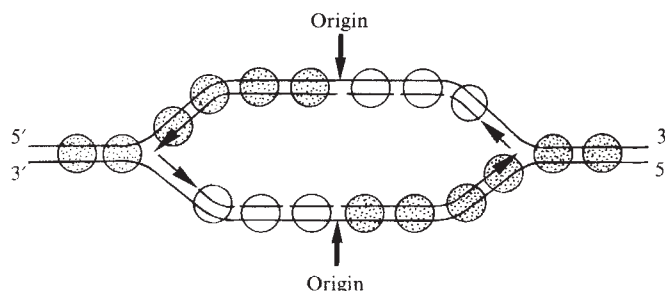


Fig. 5 A model for segregation of nucleosomes at replication forks proposed in ref. 67. Parental nucleosomes (stippled) segregate conservatively to the daughter DNA duplex containing the continuously synthesized strand, while new nucleosomes (unstippled) are assembled on the daughter duplex which is synthesized discontinuously in Okazaki fragments. Arrows indicate the direction of DNA strand growth. While this model is attractive, it is still opposed by some studies (see text).

which one of the daughter duplexes is fully assembled into nucleosomes while the other appears to be naked DNA. This suggests that parental nucleosomes pass to one of the daughter duplexes leaving the other as the site for new nucleosome formation. One of the two newly synthesized strands (the lagging strand) at a replication fork is synthesized discontinuously in Okazaki fragments (reviewed in ref. 66), and Seidman, Levine and Weintraub⁶⁷ have argued that this strand is probably the site of new nucleosome formation (Fig. 5). Although the discontinuous synthesis of DNA is not necessary to specify the spacing between adjacent nucleosomes (see above), a more attractive possibility is that the repeating nucleosome structure of chromatin may specify the periodicity of Okazaki fragments with replication proceeding by unfolding one parental nucleosome at a time⁶⁸.

Assembly of transcriptionally active nucleosomes

Digestion with nucleases indicates that transcriptionally active regions of chromatin are organized into nucleosome-like structures but that these have a different conformation from bulk chromatin (reviewed in refs 2-4). Regions of chromatin which code for RNA in a particular cell type are selectively sensitive to digestion by DNase I^{69,70} or DNase II⁷¹. The degree of DNase I sensitivity is the same whether the gene is transcribed occasionally or frequently⁷² suggesting that this altered conformation is a property of all genes which are competent to be transcribed. The boundaries of DNase I sensitive regions are precise^{54,73} and the active conformation appears to correspond to elevated levels of two non-histone proteins HMG 14 and HMG 17^{74,75} of the 'high mobility group'⁷⁶. Furthermore Weisbrod and Weintraub⁷⁵ have shown that these two proteins can restore DNase I sensitivity to salt-washed chromatin or nucleosome cores⁷⁷, indicating that the nucleosomes from which they were eluted contain an additional structural peculiarity apart from HMG 14 and 17. This could either be the presence of an additional modification or factor specifying the sites to which HMG 14 and 17 should bind, or alternatively the lack of a normal nucleosome component which competes with HMG 14 and 17 for occupation of a site within the nucleosome.

How does such a structure originate? Weintraub has shown that the DNase I sensitive structure arises within 3 min of replication of that region of DNA, suggesting that the decision to assemble potentially active or inactive chromatin occurs at the time of DNA replication⁶⁰. Therefore our understanding of the differentiation of somatic cells would be substantially enhanced by a knowledge of the mechanisms which assemble newly replicated DNA into transcriptionally competent nucleosomes and of the regulatory mechanisms which rigorously select specific regions of the genome for assembly into the active

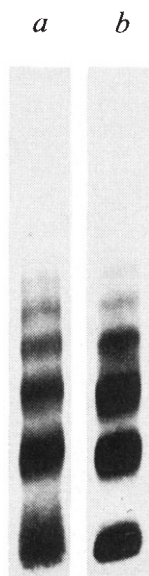


Fig. 4 Assembly of regularly spaced chains of nucleosomes by incubating ³²P DNA in homogenized eggs of *Xenopus laevis* (as in ref. 14). *a*, The chromatin assembled *in vitro* was cleaved with micrococcal nuclease deproteinized and electrophoresed in a 1.7% agarose gel. *b*, DNA fragments cleaved from rat-liver nuclei by micrococcal nuclease and ³²P-labelled by nick translation were electrophoresed on the same gel to show nucleosome spacing at 195 base pairs.

conformation. The *in vitro* nucleosome assembly systems which have been reviewed here are ideally suited for investigating these questions.

We thank C. Dingwall, J. Gurdon, R. Harland, A. Klug, J. Thomas, and S. Weisbrod for comments on the manuscript.

1. Kornberg, R. D. *A. Rev. Biochem.* **46**, 931-954 (1977).
2. Felsenfeld, G. *Nature* **271**, 115-122 (1978).
3. Chambon, P. *Cold Spring Harb. Symp. quant. Biol.* **42**, 1209-1234 (1978).
4. Lilley, D. M. J. & Pardon, J. F. *A. Rev. Genet.* **13**, 197-233 (1979).
5. Finch, J. T. *et al. Nature* **269**, 29-36 (1977).
6. Thoma, F., Koller, Th. & Klug, A. *J. Cell Biol.* **83**, 403-427 (1979).
7. Simpson, R. T. *Biochemistry* **17**, 5524-5531 (1978).
8. Mayfield, J. E., Serunian, L. A., Silver, L. M. & Elgin, S. C. R. *Cell* **14**, 539-544 (1978).
9. Elgin, S. C. R. & Weintraub, H. *A. Rev. Biochem.* **44**, 725-774 (1975).
10. Ruiz-Carrillo, A., Jorcano, J. L., Eder, G. & Lurtz, R. *Proc. natn. Acad. Sci. U.S.A.* **76**, 3284-3288 (1979).
11. Stein, A., Whitlock, J. P. & Bina, M. *Proc. natn. Acad. Sci. U.S.A.* **76**, 5000-5004 (1979).
12. Laskey, R. A., Honda, B. M., Mills, A. D. & Finch, J. T. *Nature* **275**, 416-420 (1978).
13. Germond, J. E., Rouviere-Yaniv, J., Yaniv, M. & Brutlag, D. *Proc. natn. Acad. Sci. U.S.A.* **76**, 3779-3783 (1979).
14. Laskey, R. A., Mills, A. D. & Morris, N. R. *Cell* **10**, 237-243 (1977).
15. Baldi, M. I., Mattoccia, E. & Tocchini-Valentini, G. P. *Proc. natn. Acad. Sci. U.S.A.* **75**, 4873-4876 (1978).
16. Nelson, T., Hsieh, T.-S. & Brutlag, D. *Proc. natn. Acad. Sci. U.S.A.* **76**, 5510-5514 (1979).
17. Axel, R., Melchior, W., Sollner-Webb, B. & Felsenfeld, G. *Proc. natn. Acad. Sci. U.S.A.* **71**, 4101-4105 (1974).
18. Oudet, P., Gross-Bellard, M. & Chambon, P. *Cell* **4**, 281-300 (1975).
19. Germond, J. E., Hirt, B., Oudet, P., Gross-Bellard, M. & Chambon, P. *Proc. natn. Acad. Sci. U.S.A.* **72**, 1843-1847 (1975).
20. Camerini-Otero, R. D., Sollner-Webb, B. & Felsenfeld, G. *Cell* **8**, 333-347 (1976).
21. Steinmetz, M., Streeck, R. E. & Zachau, H. G. *Eur. J. Biochem.* **83**, 615-628 (1978).
22. Champoux, J. J. *A. Rev. Biochem.* **47**, 449-479 (1978).
23. Laskey, R. A., Honda, B. M., Mills, A. D., Harland, R. M. & Earnshaw, W. C. in *Eukaryotic Gene Regulation, ICN-UCLA Symposium 14* (eds Maniatis, T., Axel, R. & Fox, C. F. 551-559) (Academic, New York, 1979).
24. Krohne, G. & Franke, W. *Proc. natn. Acad. Sci. U.S.A.* **77**, 1034-1038 (1980).
25. Krohne, G. & Franke, W. *Expl. Cell Res.* (in the press).
26. Mills, A. D., Laskey, R. A., Black, P. & De Robertis, E. M. *J. molec. Biol.* **139**, 561-568 (1980).
27. Earnshaw, W. C., Honda, B. M., Laskey, R. A. & Thomas, J. O. *Cell* (in the press).
28. Voordouw, G. & Eisenberg, H. *Nature* **273**, 446-448 (1978).
29. Stein, A. *J. molec. Biol.* **130**, 103-104.
30. Wyllie, A. H., Laskey, R. A., Finch, J. & Gurdon, J. B. *Dev. Biol.* **64**, 178-188 (1978).
31. Adamson, E. D. & Woodland, H. R. *J. molec. Biol.* **88**, 263-285 (1974).
32. Woodland, H. R. & Adamson, E. D. *Dev. Biol.* **57**, 118-135 (1977).
33. Worcel, A., Han, S. & Wong, M. L. *Cell* **15**, 969-977 (1978).
34. Senshu, T., Fukada, M. & Ohashi, M. *J. Biochem. (Japan)* **84**, 985-988 (1978).
35. Cremisi, C., Chestier, A. & Yaniv, M. *Cell* **12**, 947-951 (1977).
36. Rabinowitz, M. *J. Morphol.* **69**, 1-49 (1941).
37. Blumenthal, A. B., Kriegstein, H. J. & Hogness, D. S. *Cold Spring Harb. Symp. quant. Biol.* **38**, 205-233 (1973).
38. Louie, A. J. & Dixon, G. H. *Proc. natn. Acad. Sci. U.S.A.* **69**, 1975-1979 (1972).
39. Ruiz-Carrillo, A., Wang, L. J. & Allfrey, V. *Science* **190**, 117-128 (1975).
40. Jackson, V., Shires, A., Tanphaichit, N. & Chalkley, R. *J. molec. Biol.* **104**, 471-483 (1976).
41. Hewish, D. R. & Burgoyne, L. A. *Biochem. biophys. Res. Commun.* **52**, 504-510 (1973).
42. Noll, M., *Nature* **251**, 249-251 (1974).
43. Morris, N. R., *Cell* **9**, 627-632 (1976).
44. Noll, M., *Cell* **8**, 349-355 (1976).
45. Thomas, J. O. & Thompson, R. J. *Cell* **10**, 633-640 (1977).
46. Compton, J. L., Bellard, M. & Chambon, P. *Proc. natn. Acad. Sci. U.S.A.* **73**, 4382-4386 (1976).
47. Spadafora, C., Bellard, M., Compton, J. L. & Chambon, P. *FEBS Lett.* **69**, 281-285 (1976).
48. Woodland, H. R. *Dev. Biol.* **68**, 360-370 (1979).
49. Ponder, B. A. J. & Crawford, L. V. *Cell* **11**, 35-49 (1977).
50. Varshavsky, A. J., Sundin, O. & Bohn, M. *Cell* **16**, 453-466 (1979).
51. Waldeck, W., Föhrling, B., Chowdhury, K. Gruss, P. & Sauer, G., *Proc. natn. Acad. Sci. U.S.A.* **75**, 5964-5968 (1978).
52. Scott, W. A. & Wigmore, D. J. *Cell* **15**, 1511-1518 (1978).
53. Wittig, B. & Wittig, S. *Cell* **18**, 1173-1183 (1979).
54. Wu, C., Bingham, P. M., Livak, K. J., Holmgren, R. & Elgin, S. C. R. *Cell* **16**, 797-806 (1979).
55. Musich, P. R., Maio, J. J. & Brown, F. L. *J. molec. Biol.* **117**, 637-655 (1977).
56. Seale, R. L. *Nature* **225**, 247-249 (1975).
57. Seale, R. L. *Cold Spring Harb. Symp. quant. Biol.* **42**, 433-438 (1978).
58. McKnight, S. L. & Miller, O. L. Jr *Cell* **12**, 795-804 (1977).
59. Levy, A. & Jakob, K. M. *Cell* **14**, 259-267 (1978).
60. Weintraub, H. *Nucleic Acids Res.* **7**, 781-792 (1979).
61. Leffak, M., Grainger, R. & Weintraub, H. *Cell* **12**, 837-845 (1977).
62. Seale, R. L. *Cell* **9**, 423-429 (1976).
63. Hancock, R. *Proc. natn. Acad. Sci. U.S.A.* **75**, 2130-2134 (1978).
64. Jackson, V., Granner, D. & Chalkley, R. *Proc. natn. Acad. Sci. U.S.A.* **73**, 2266-2269 (1976).
65. Riley, D. & Weintraub, H. *Proc. natn. Acad. Sci. U.S.A.* **76**, 328-332 (1979).
66. Hand, R. *Cell* **15**, 317-325 (1978).
67. Seidman, M. M., Levine, A. J. & Weintraub, H. *Cell* **18**, 439-449 (1979).
68. Kornberg, R. D. *Science* **184**, 868-871 (1974).
69. Garel, A. & Axel, R., *Proc. natn. Acad. Sci. U.S.A.* **73**, 3966-3970 (1976).
70. Weintraub, H. & Groudine, M. *Science* **93**, 848-853 (1976).
71. Gottesfeld, J. M., Bagi, G., Berg, B. & Bonner, J. *Biochemistry* **15**, 2472-2483 (1976).
72. Garel, A., Zolan, M. & Axel, R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 4867-4871 (1977).
73. Flint, S. J. & Weintraub, H. M. *Cell* **12**, 783-794 (1977).
74. Levy, W. B. & Dixon, G. *Nucleic Acids Res.* **5**, 4155-4163 (1978).
75. Weisbrod, S. & Weintraub, H. *Proc. natn. Acad. Sci. U.S.A.* **76**, 630-634 (1979).
76. Goodwin, G. H., Walker, J. M. & Johns, E. W. in *The Cell Nucleus* Vol. 6 (ed. Busch, H.) pp. 181-219 (1978).
77. Weisbrod, S., Groudine, M. & Weintraub, H. *Cell* **19**, 289-301 (1980).

ARTICLES

An Archaean sub-seafloor geothermal system, 'calc-alkali' trends, and massive sulphide genesis

P. J. MacGeehan* & W. H. MacLean

Abitibi Project, Department of Geological Sciences, McGill University, 3450 University Street, Montreal, Quebec, Canada H3A 2A7

Archaean Noranda-type exhalative massive sulphide ores were formed by sub-seafloor geothermal activity which leached mafic constituents and ore metals from iron-rich bimodal (basalt-rhyolite) tholeiites and converted them to hydrothermally altered greenschist facies 'intermediate' rocks of apparent calc-alkaline affinity.

A MAJOR problem in the study of Archaean and younger greenstone belts is the identification of the primary geochemistry of volcanic rocks that have been subjected to extensive rock-water interaction. The rocks, mostly fine grained basalts exhibiting greenschist facies mineralogy, are normally classified on the basis of chemical composition and normative mineralogy on the assumption that rock-water interaction was largely an isochemical process. This hypothesis apparently succeeds in large parts of the greenstone belts where the volcanic rocks were only weakly hydrated during halmyrolysis and burial metamor-

phism, but, in some areas, sub-seafloor geothermal activity contemporaneous with volcanism produced more intense rock-water interaction with higher water-to-rock ratios. It is in these areas of intense hydrothermal alteration, seafloor exhalative activity and associated massive sulphide mineralization, that the primary petrological and chemical identity of the volcanic rocks has been radically changed, causing them to be incorrectly identified using standard geochemical rock classification schemes.

The nature of this problem became evident during research in the Matagami, Quebec Archaean volcanic centre and mining district in the Abitibi greenstone belt, where sulphide ores are associated with several thick domal accumulations of rhyolite

* Present address: Department of Geological Sciences, Queen's University, Kingston, Ontario, Canada K7L 3N6.

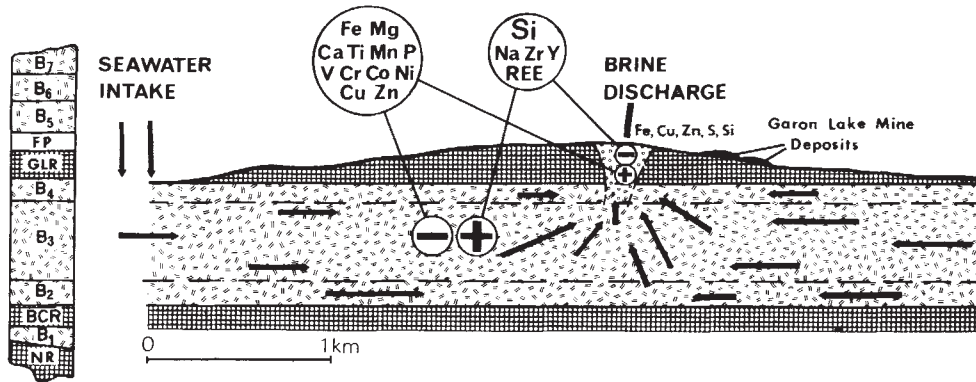


Fig. 1 Volcanic stratigraphy and diagrammatic vertical section through the Garon Lake geothermal system. B₁–B₇, basalts; FP, andesite; GLR, Garon Lake rhyolite; BCR, Bell Channel rhyolite; NR, Norita rhyolite. The pattern with circles denotes intensely chloritized Garon Lake rhyolite.

interbedded with iron-rich pillowed basalt flows. These rocks were formerly described as 'calc-alkaline'¹ owing to the predominance of rocks of 'intermediate' chemical composition (andesite–dacite–rhyolite). These show a large scatter rather than a systematic alkali-enrichment on an AFM plot, a feature characteristic of many mineralized Archaean volcanic centres^{2–5}. However, they are now known to constitute a hydrothermally altered bimodal basalt–rhyolite suite of clear tholeiitic affinity⁶. Their apparent 'calc-alkaline' traits reflect a hydrothermal overprint caused by rock–water interaction during sub-sea floor geothermal activity as a modified seawater brine intensely spilitized, silicified and leached metals from large volumes of iron-rich tholeiitic basalt before discharging through overlying rhyolite to deposit massive sulphides on the sea floor.

In this article we examine the petrography and geochemistry of the rock–water interaction in the Garon Lake geothermal system at Matagami, and compare and contrast the trends produced with recognized igneous fractionation trends.

The Garon Lake geothermal system

The Garon Lake geothermal system developed within a 500-m thick sequence of tholeiitic pillowed basalt (units B₂–B₄) after extrusion of the Garon Lake rhyolite (Fig. 1). The basalts are underlain by the Bell Channel rhyolite, thus the reservoir rocks of the system are confined to a narrow, well-defined stratigraphic interval in which the individual flow-units have been mapped in detail and followed along strike for several kilometres. The basalts are intruded by contemporaneous gabbro sills of very similar mineralogical and chemical composition, several of which may have been lateral feeders to the pillowed

flows. The geological setting, petrology and primary geochemistry of the units have been discussed elsewhere^{6,7}, where they were shown to form part of a bimodal basalt–rhyolite suite of tholeiitic affinity.

Most of the volcanic rocks have been hydrothermally altered to some degree: the basalts are intensely spilitized and silicified, whereas the domal parts of the Garon Lake rhyolite are pervasively chloritized and transected by zones of disseminated sulphide mineralization. Earlier studies have shown that this alteration accompanied sub-seafloor geothermal activity which began while the volcanic rocks and contemporaneous intrusions were still hot, and terminated before the extrusion of volcanic rocks overlying the Garon Lake rhyolite^{7,8}. At the time when the geothermal system was active, seawater was convected through basalt from intake zones at the periphery of the rhyolite and then discharged through permeable zones (now alteration pipes) in the rhyolite dome. The exhalative massive sulphide lenses at the Garon Lake mine formed down slope from the western alteration pipe⁹. The geothermal system had a lateral extent > 5 km (Fig. 1), indicating that > 10 km³ of volcanic rock were subjected to hydrothermal alteration.

Rock–water interaction

The reservoir rocks are bleached and mottled pillowed basalts that would almost invariably be called 'andesite' in the field. Two main stages of hydrothermal alteration can be recognized petrographically: (1) an early and minor stage of dominantly destructive (argillic) alteration that accompanied localized convective fluid movement in water-bearing basalt bordering some of the contemporaneous gabbro sills⁶; and (2) a later and

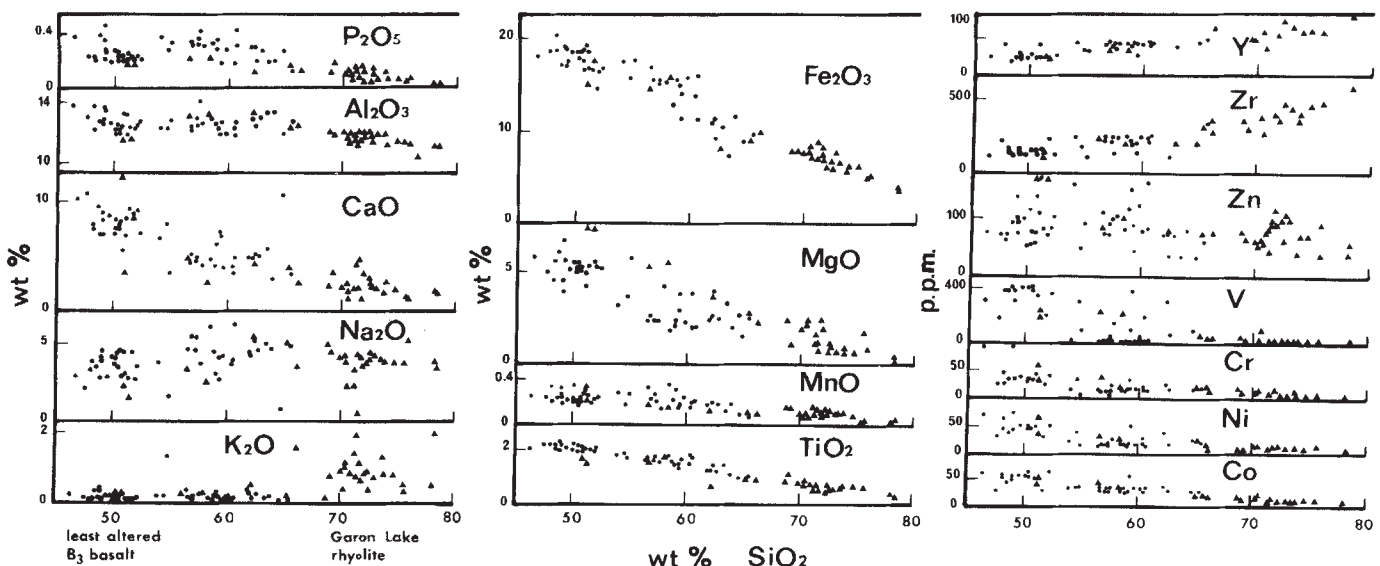


Fig. 2 Plot of major and trace elements against SiO₂ for the B₃ basalt and Garon Lake rhyolite units. Alteration produces the continuum of data. ●, B₃ basalt samples; ▲, Garon Lake rhyolite samples. Total iron as Fe₂O₃.

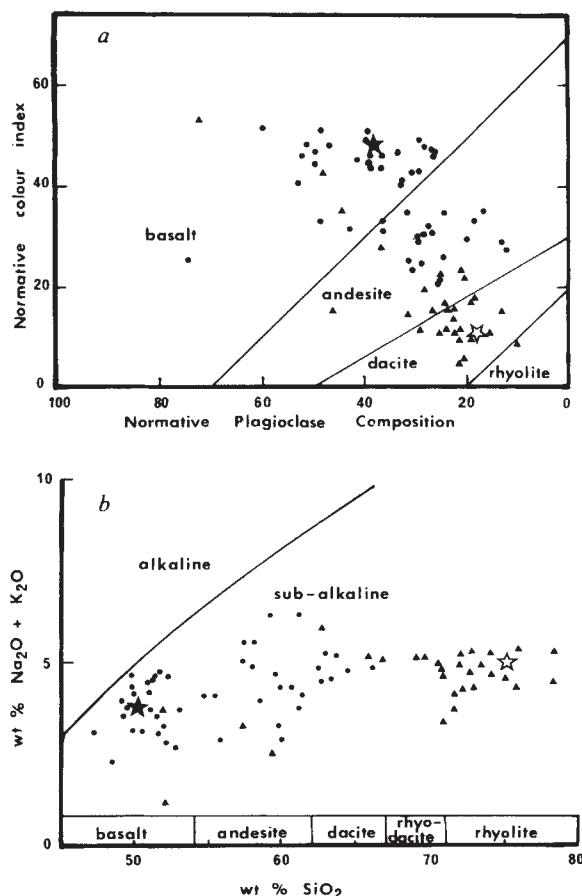


Fig. 3 *a*, Normative CI against normative An¹⁵. *b*, Total alkalis against SiO₂; the rock classification based on SiO₂ content is after Gélinas *et al.*³, although those by Spitz and Darling⁴ and Jolly¹⁶ are very similar. The mean composition of least altered samples are: ★, B₃ basalt; and ☆, Garon Lake rhyolite.

widespread stage of intense spilitization that accompanied large-scale convective circulation of fluid after the reservoir rocks were sealed from above by the Garon Lake rhyolite (Fig. 1).

Spilitization is localized around pervasive quartz-epidote-filled veinlets and cylindrical pipes (alteration patches¹⁰) which transect the basalt section. These structures form an interconnecting network interpreted as conduits of concentrated fluid flow at the time the geothermal system was active. Whereas the smaller veinlets have a random orientation, the larger veinlets and pipes are invariably orientated parallel to stratigraphy, indicating a dominantly horizontal vector of fluid movement within the reservoir rocks (Fig. 1).

The basalt around the quartz-epidote structures has been intensely bleached and silicified to light grey and grey-green domains up to one or several metres in diameter. Although the igneous fabric was almost entirely preserved in these domains the mineral assemblages were drastically changed: plagioclase phenocrysts and laths were replaced by single crystals of inclusion-free albite that commonly retain original twin planes, pyroxene was replaced by actinolite (twin planes, exsolution lamellae and subophitic textures partly preserved), and ulvospinel first crystallized ilmenite in trelliswork patterns but was then partly replaced by actinolite and sphene. With advanced spilitization, epidote, magnetite, ilmenite and sphene were replaced by quartz, albite, minor chlorite and rutile, the latter as needles preserving the outline of former ilmenite grains. This intense spilitization grades outward to a zone of weak but pervasive hydration where basalt has a uniformly dark green colour and a more or less constant chemical composition. Here, the pyroxene is pseudomorphed by dark green actinolite,

plagioclase is saussuritized, and the groundmass consists of ulvospinel granules and sodic plagioclase crystallites in a chloritic, devitrified glass matrix.

Overall, the mineralogical changes in the basalts were accompanied by significant decreases in colour index and increases in hardness so that they now frequently resemble andesites and dacites in the field. The role of quartz is particularly important. While not a primary phase or a product of isochemical devitrification of basaltic glass, quartz increases significantly with increasing degree of alteration.

The Garon Lake rhyolite (dacite in some classification schemes^{6,11,15}) is devitrified quartzo-feldspathic and generally spherulitic acid glass that contains up to 20% inverted β -quartz and sodic plagioclase phenocrysts and rare granophyre fragments⁶⁻⁸. Zones of intense chloritization that constitute the alteration pipes are mineralized with chalcopyrite and pyrite (Fig. 1). The direct replacement of rhyolite glass (~75% SiO₂) by massive chlorite (~35% SiO₂) proceeded without obvious volume change; β -quartz phenocrysts were corroded and plagioclase grains were either partly corroded or entirely replaced by granular quartz plus chlorite or massive chlorite.

Geochemical changes

Intense spilitization of the reservoir rocks was accompanied by major chemical adjustments, the most important being the significantly higher SiO₂ content of altered samples. Unaltered or least altered samples of the chilled rims and cores

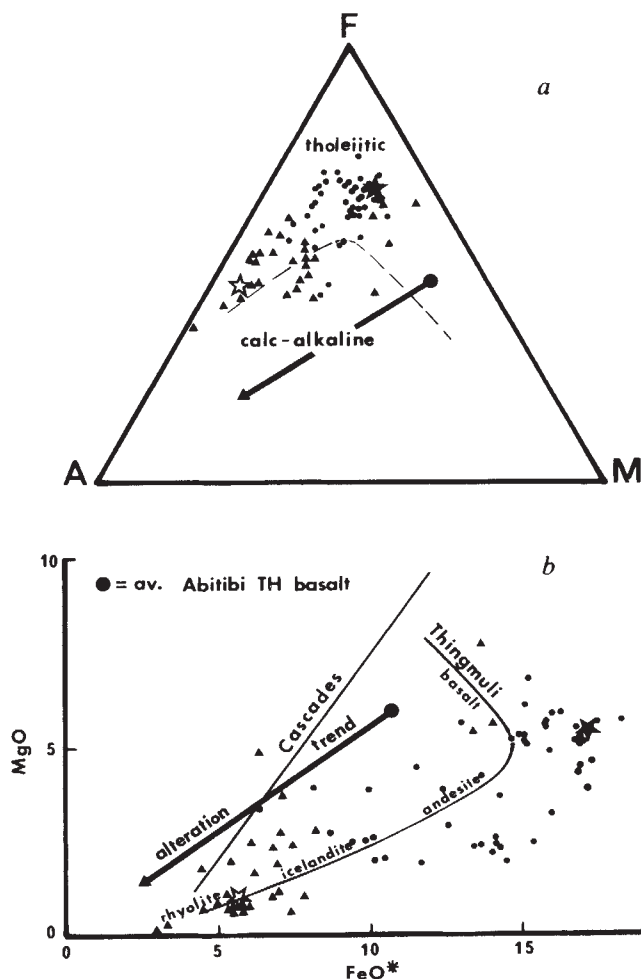


Fig. 4 *a*, AFM plot. The dividing line is after Irvine and Baragar¹⁵. *b*, MgO against FeO* (total iron as FeO). The Thingmuli¹⁷ (tholeiitic) and Cascades (calc-alkaline) trends are shown. A plot, ●, of the mean composition of basalts from the Abitibi greenstone belt¹⁸ shows how alteration would move them into the calc-alkaline field.

of pillows lie in the compositional range 48–52% SiO_2 , but hydrothermally altered samples show a progressive increase in SiO_2 to 65% in the B_3 unit and up to 75% in B_4 (ref. 8). Increases in SiO_2 in B_3 (Fig. 2) are accompanied by: (1) a linear depletion in FeO^* , MgO , CaO , TiO_2 , P_2O_5 , V , Cr , Mn , Co , Ni , Cu , Zn and Cd , reflecting their dissolution during breakdown and reconstitution of the mafic minerals and plagioclase; (2) a significant increase in Na accompanying the albitization of plagioclase; and (3) comparable increases in Zr , Y and REE by factors of up to three. K_2O , which probably had a very low initial content, remains low in the altered rocks. Although scatter is considerable, the ratios of depleted elements (FeO^*/MgO , $\text{FeO}^*/\text{TiO}_2$, and so on) remained essentially constant and identical to the igneous ratios, indicating that their dissolution was a congruent, non-selective process. Hydrothermal alteration in the remaining basalt flows (units B_2 and B_4) was very similar, although changes in the intensity of silicification, in the $\text{Na}_2\text{O}/\text{SiO}_2$ enrichment, and in the behaviour of P_2O_5 and Al_2O_3 did occur with stratigraphic height within the geothermal system^{7,8}.

Two lines of evidence, the petrographic observation of the pseudomorphic replacement of the igneous rock fabric and the immobile behaviour of Al_2O_3 (Fig. 2), indicate that rock–water interaction was essentially a constant volume process. The incremental enrichment in SiO_2 and Na_2O was therefore accompanied by the active congruent dissolution and transport of FeO^* , MgO , CaO , TiO_2 and the trace elements listed above, because only SiO_2 , CaO and Al_2O_3 (quartz and epidote) were later precipitated in the hydrothermal conduits¹⁰. This alteration accompanied an increase in the temperature of the fluid. The precipitation of silica in the central part of the geothermal system occurred around the critical point of the brine ($T \sim 400^\circ\text{C}$; $P < 0.75$ kbar)^{12–14}, where SiO_2 solubility decreases rapidly with rising temperatures while that of most other components continues to increase. Similarly, as the metal-charged brine rose through the rhyolite to the sea floor it cooled, leached silica from rhyolite, and precipitated most of the elements dissolved from the underlying reservoir rocks as minerals in the alteration pipe and massive sulphides on the sea floor (Fig. 2).

Significance of the alteration trends

Rock–water interaction during sub-seafloor geothermal activity caused the congruent dissolution of many major and trace elements from basalt, their transport by fluid, and their later precipitation in similar ratios in rhyolite approaching the discharge points of the system. Cumulatively, this resulted in the compositional gap of the bimodal basalt–rhyolite suite being spanned by a continuum of altered rock compositions (Fig. 2). Using as an example only one basalt flow from the reservoir and the overlying Garon Lake rhyolite (Figs 2–6), the geochemistry of the hydrothermally altered rocks seems to be indistinguish-

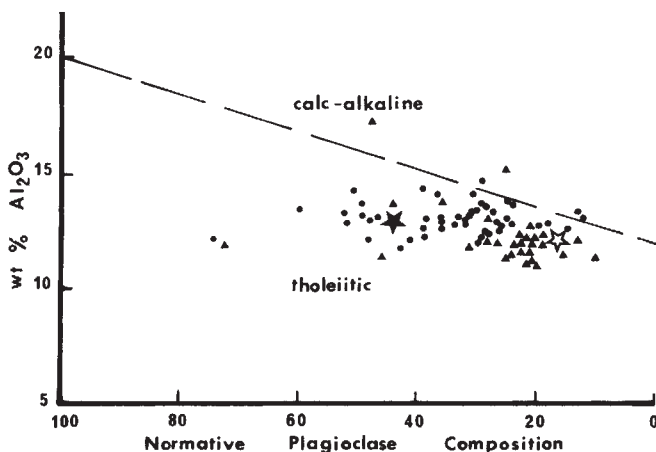


Fig. 5 Comparison of Al_2O_3 with normative An^{15} .

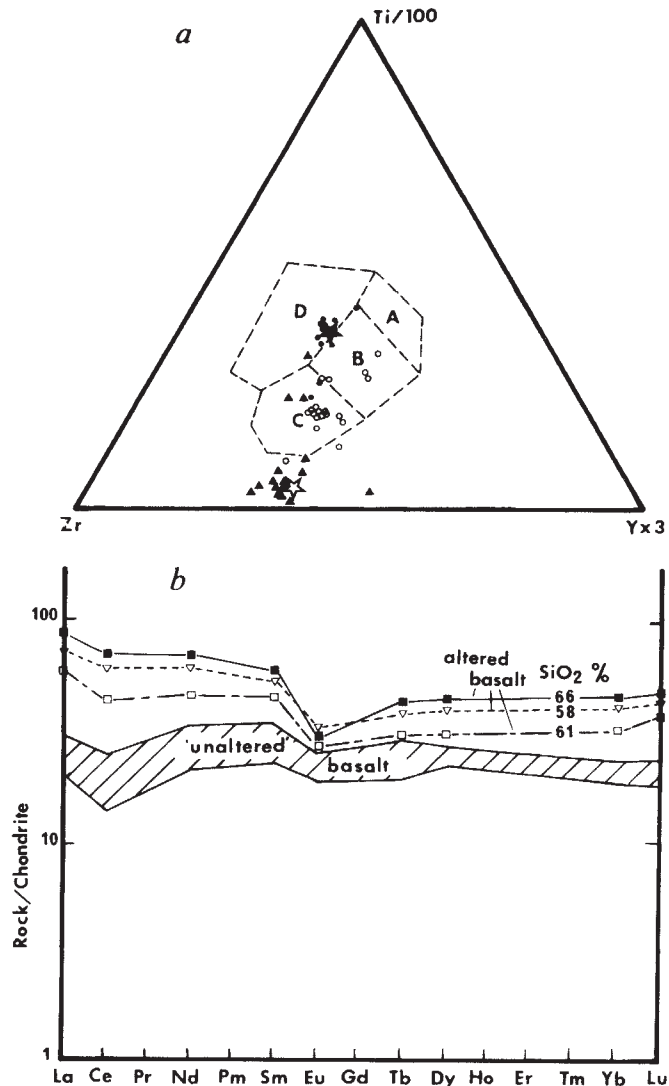


Fig. 6 a, Ti–Zr–Y plot¹⁹. ●, B_3 basalt with $< 54\%$ SiO_2 ; ○, $> 54\%$ SiO_2 ; ▲, Garon Lake rhyolite. Low-K tholeiites occupy fields A and B; ocean floor basalts, B; calc-alkaline basalts, B and C; within-plate basalts, D. b, Rare-earth element composition of B_3 basalt samples. SiO_2 contents of altered samples are indicated. REE values are normalized to the Leedy chondrite²⁴. Analyses by neutron activation at X-ray Assay Laboratories, Toronto, Canada.

able from the spectrum of intermediate igneous rock compositions. Hydrothermally altered basalt and rhyolite samples span the field through andesite and dacite on a normative colour index versus normative plagioclase composition diagram (Fig. 3a) used to classify volcanic rock types¹⁵, and classification systems based on SiO_2 content^{3,4,16} (see Fig. 3b).

Furthermore, as rock–water interaction involved the net substitution of salic (SiO_2 , Na_2O , and so on) for basic (FeO^* , MgO , CaO , TiO_2 , and so on) components, the resulting chemical alteration patterns resemble igneous fractionation trends. For example, most major and trace elements show a variation typical of volcanic rock suites when plotted against a fractionation index (for example, SiO_2 , Fig. 2; differentiation index and solidification index plots show similar patterns). However, despite Na metasomatism during spilitization of basalt, altered samples remain in the sub-alkaline field on an alkali–silica plot (Fig. 3b). The resemblance with a fractionation trend is also seen on an AFM plot, widely used to distinguish tholeiitic from calc-alkaline rocks; the least altered basalt samples lie close to the FM join and altered samples extend towards the alkali corner (Fig. 4a). The resemblance is maintained in plots not involving the alkalis, for example, FeO^* versus MgO (Fig. 4b),

where, despite the scatter, the alteration parallels the later stages of the Thingmuli tholeiitic trend¹⁷. There is no evidence of iron enrichment, however, and the resemblance to a tholeiitic trend is dictated by the high initial FeO*/MgO of this basalt which precludes a calc-alkaline association. This would not have been so in most Archaean centres where the basalts generally have much lower FeO*/MgO ratios (see Fig. 4 for plots of average Abitibi basalt¹⁸). In its essential characteristics (alkali enrichment with depletion of mafic oxides) the alteration mimics a calc-alkaline fractionation trend, and would easily be confused for one in many mineralized volcanic centres. This apparent calc-alkaline affinity is indicated on many other major element plots except that of Al₂O₃ against normative plagioclase composition (Fig. 5) where the low mobility of Al₂O₃ during alteration confines the data to the tholeiitic field.

The whole trace element suite was also mobile during alteration: V, Cr, Co, Ni, Cu, Zn and Cd decreased, and Zr, Y and REE increased with increasing SiO₂ content (Figs 2–6). A plot of the 'immobile' elements Ti, Zr and Y (Fig. 6a) shows how mobile they have been. The negative correlation of Ti for Zr and Y that can be seen in Fig. 2 produces a continuum of points between the original basalt and rhyolite on Fig. 6a. This diagram has been used successfully to distinguish the tectonic setting of modern basalts (<54% SiO₂)¹⁹, and the correlations made there have been frequently used to classify older rocks^{20–23}. Most samples from the reservoir rocks of the geothermal system with less than 54% SiO₂ do plot in a small area of Fig. 6a, but the tendency of the altered rocks to move to the calc-alkaline field is apparent. Plots with Cr, Co and Ni yield similar results. The REE are also mobile but retain the flat profile (Fig. 6b) typical of fractionated tholeiites, lacking the characteristic strong LREE enrichment and high La/Yb ratios of calc-alkaline rocks. Rather

than being immobile, these elements actually provide a good monitor of progressive alteration in a single flow.

Conclusions

Intense rock–water interaction in the Garon Lake sub-seafloor geothermal system accompanied spilitization of volcanic rocks and contemporaneous high-level gabbro sills. Massive chemical exchange occurred in and adjacent to zones of dynamic flow of seawater-derived brine; essentially isochemical hydration occurred where flow was sluggish or stagnant. Hydrothermal alteration preserved volcanic features (pillow selvages, amygdulites, hyaloclastite-filled interstices) and pseudomorphed igneous fabric as it converted bimodal basalt–rhyolite tholeiites to rocks that are petrographically indistinguishable from metamorphosed andesites and dacites. The major and trace element geochemistry of the altered rocks are 'intermediate' in character and show calc-alkaline affinity on most discrimination diagrams.

The predominance of these 'intermediate' rocks at Matagami and in other mineralized volcanic centres in the Archaean of Canada, led to the original assumption that exhalative massive sulphide ores of the Noranda type were associated with 'calc-alkaline' volcanic rocks, and, by inference, with active volcanism along consuming plate boundaries^{1,4,25}. It is now evident that these conclusions may be in error. Rather than denoting a tectonic setting, these 'intermediate' 'calc-alkaline' volcanic rocks may in fact reflect the geochemical imprint of sub-seafloor geothermal activity that generated important ore deposits.

This work was supported by a Quebec Government FCAC grant 649 and NSERC grant A7719. We thank R. F. Martin, A. J. Hynes, G. R. Webber, D. Francis, K. St. Seymour, C. J. Hodgson and T. H. Pearce for helpful comments.

Received 8 May; accepted 5 June 1980.

1. Descarreaux, J. *CIM Bull.* **6**, 61–69 (1973).
2. Allard, G. O. *Quebec Dept Nat. Resour., Miner. Dep. Rep. DP-368* (1976).
3. Gélinas, L. et al. in *Volcanic Regimes in Canada* (eds Baragar, W. R. A., Coleman, L. C. & Hall, J. M.) Spec. Pap. 16, 265–295 (Geological Association of Canada, 1977).
4. Spitz, G. & Darling, R. *Can. J. Earth Sci.* **12**, 1820–1849 (1975).
5. Beswick, A. E. & Soucie, G. *Precamb. Res.* **6**, 235–248 (1978).
6. MacGeehan, P. J. & MacLean, W. H. *Nature* **283**, 153–157 (1980).
7. MacGeehan, P. J. thesis, McGill Univ., Montreal (1979).
8. MacGeehan, P. J. *Can. J. Earth Sci.* **15**, 551–570 (1978).
9. MacLean, W. H. & MacGeehan, P. J. *IREM/MERI Case Hist.* 76–1, (McGill University, Montreal, 1976).
10. Harrigan, D. B. & MacLean, W. H. *Can. J. Earth Sci.* **13**, 500–511 (1976).
11. Streckeisen, A. *Geology* **7**, 331–335 (1979).
12. Kennedy, G. C. *Econ. Geol.* **45**, 629–653 (1950).
13. Sourirajan, S. & Kennedy, G. C. *Am. J. Sci.* **260**, 115–141 (1962).
14. Spooner, E. T. C. & Fyfe, W. S. *Contr. Miner. Petrol.* **42**, 287–304 (1973).
15. Irvine, T. N. & Baragar, W. R. A. *Can. J. Earth Sci.* **8**, 523–548 (1971).
16. Jolly, W. T. in *Volcanic Regimes in Canada* (eds Baragar, W. R. A., Coleman, L. C. & Hall, J. M.) Spec. Pap. 16, 311–330 (Geological Association of Canada, 1977).
17. Carmichael, I. S. E. *J. Petrol.* **5**, 435–460 (1964).
18. Goodwin, A. M. in *Volcanic Regimes in Canada* (eds Baragar, W. R. A., Coleman, L. C. & Hall, J. M.) Spec. Pap. 16, 205–241 (Geological Association of Canada, 1977).
19. Pearce, J. A. & Cann, J. R. *Earth planet. Sci. Lett.* **19**, 290–300 (1973).
20. Morrison, M. A. *Earth planet. Sci. Lett.* **39**, 407–416 (1978).
21. Davies, J. F., Grant, R. W. E. & Whitehead, R. E. S. *Can. J. Earth Sci.* **16**, 305–311 (1979).
22. Winchester, J. A. & Floyd, P. A. *Earth planet. Sci. Lett.* **28**, 459–469 (1976).
23. Smith, R. E. & Smith, S. E. *Earth planet. Sci. Lett.* **32**, 114–120 (1976).
24. Nakamura, N. *Geochim. cosmochim. Acta* **38**, 757–775 (1974).
25. Sangster, D. F. & Scott, S. D. in *Handbook of Strata-Bound and Stratiform Ore Deposits* Vol. 6 (ed. Wolf, K. H.) 129–222 (Elsevier, Amsterdam, 1976).

Antigenic drift between the haemagglutinin of the Hong Kong influenza strains A/Aichi/2/68 and A/Victoria/3/75

Martine Verhoeyen, Rongxiang Fang, Willy Min Jou, René Devos, Danny Huylebroeck, Eric Saman & Walter Fiers

Laboratory of Molecular Biology, State University of Ghent, Ledeganckstraat 35, B-9000 Ghent, Belgium

A DNA copy of the gene coding for the influenza A/Aichi/2/68 haemagglutinin protein was cloned in the plasmid pBR322 and the complete nucleotide sequence determined. Comparison of this primary structure and the deduced amino acid sequence with the haemagglutinin gene and protein of strains belonging to the same (H3) subtype and to different subtypes, of both human (H2) and avian (Hav1) origin, documents further at the molecular level the two independent modes of antigenic variation of the virus—drift and shift.

THE ability of type A influenza strains to change their antigenic properties is a major obstacle to controlling the viral disease by vaccination, and the virus continues to be an important cause of

illness in the human population. Two types of change have been recognized: antigenic 'drift' seems to be the result of a certain number of local alterations in the viral surface proteins (thought

to be selected under the pressure of the immune system), whereas antigenic 'shift' is a radical change of the antigenic properties of the surface protein, leading to the emergence of new pandemic strains. The origin of these pandemic strains is unknown, but it has been speculated that they result from a recombination event between a human virus and an influenza strain from an animal reservoir¹.

Type A influenza viruses have a negative-stranded RNA genome consisting of eight distinct RNA segments²⁻⁴. The two surface antigens, haemagglutinin and neuraminidase, are glycoproteins and have been assigned to viral RNAs 4 and 5 or 6 (depending on the strain), respectively^{2,3}. We have recently determined the primary structure of the haemagglutinin gene from the relatively recent human influenza strain A/Victoria/3/75 (H3N2), starting from cloned DNA⁵ and using fast DNA sequencing techniques. Compared with RNA and protein sequencing methods, this approach has the advantage of much greater speed and accuracy and avoids special problems, such as those encountered during the sequencing of hydrophobic regions of polypeptides.

We describe here the cloning and elucidation of the structure of the haemagglutinin gene of A/Aichi/2/68, which formed the beginning of the so-called Hong Kong (H3N2) period. The gene sequence and deduced amino acid sequence are compared with those of our previously reported Victoria strain⁵, with an intermediate H3N2 isolate, A/Memphis/102/72 (this haemagglutinin was sequenced as cloned DNA⁶ and except for a small hydrophobic region also as protein^{7,8}), with amino acid data from monoclonal variants⁹ and field strains¹⁰, with the gene sequence of A/Japan/305/57 (H2N2)¹¹, which occurred at the start of the 1957-67 influenza wave, and with the fowl plague haemagglutinin gene (Hav1)¹².

Cloning the A/Aichi/2/68 haemagglutinin gene

The high yielder, recombinant strain X31, which carries the haemagglutinin and neuraminidase genes of A/Aichi/2/68, was used as the RNA source in our study. Double-stranded (ds) DNA copies of the haemagglutinin genetic information were synthesized and cloned essentially as described previously for the haemagglutinin gene of the A/Victoria/3/75 strain⁵. The major steps of the procedure are briefly summarized as follows: the mixture of eight negative-stranded X31 virion RNA segments was polyadenylated at the 3' end¹³, converted into dsDNA copies, sized on a 1.5% agarose gel and the material corresponding to full-length haemagglutinin dsDNA recovered from *Escherichia coli* K12 HB101 (refs 5, 14). Thirty-nine dTTP and annealed to oligo(dA)-tailed plasmid pBR322 in the *Pst*I restriction site, and the resulting hybrid was used to transform *Escherichia coli* K12 HB101 (refs 5, 14). Thirty-nine tetracycline-resistant colonies were obtained. When hybridized¹⁵ to a ³²P-labelled gene 4 RNA probe, four of the clones showed positive hybridization. One plasmid (pXHA14) containing an insert of approximately 1,860 base pairs (including the A·T tails used for cloning) was chosen for further characterization.

Nucleotide sequence of pXHA14

The insert in plasmid pXHA14 was characterized by restriction mapping according to Smith and Birnstiel¹⁶. Subsequently, fragments were prepared by digestion with the appropriate restriction enzymes, 5'-terminally labelled, and sequenced according to the method of Maxam and Gilbert¹⁷. The nucleotide sequence is shown in Fig. 1 as the coding information or complement of the viral RNA. As the structure is highly homologous to the previously determined Victoria haemagglutinin gene (Fig. 1), we did not systematically attempt to read both strands. Nevertheless, more than half of the sequence was confirmed by analysis of both strands.

As expected from the cloning strategy used^{5,14,18} (initiating second DNA strand synthesis starting from a terminal hairpin of

the first DNA strand, followed by S₁ nuclease treatment), a small portion of the haemagglutinin genetic information corresponding to the 3' end of the cDNA is missing from the clone. As we started from negative-stranded influenza RNA, this region corresponds to the 3' end of the coding information (beyond the arrow in Fig. 1). Because two other strains belonging to the same subtype have been shown to share exactly the same 3'-non-coding sequence^{5,6}, we did not attempt to complete this sequence for the Aichi strain. Assuming identity in this region, 27 nucleotides would be missing from the pXHA14 plasmid, compared with 24 nucleotides missing from plasmid pVHA14, which contains the Victoria haemagglutinin information⁵. Taking into account an additional 27 base pairs at the 3' end of the influenza-specific information in pXHA14, the total length of the Aichi haemagglutinin RNA is 1,765 nucleotides: 29 nucleotides precede the signal for the AUG translation initiation codon, the translated region extends over 1,698 nucleotides (specifying 566 amino acids), and the (presumed) 3'-terminal noncoding stretch comprises 38 nucleotides (including the signal for the UGA terminator).

The detailed organization of the haemagglutinin RNAs belonging to the Hong Kong (H3) subtype (Victoria⁵, Memphis⁶⁻⁸, Aichi) seems to be identical except for the insertion of three nucleotides (specifying an asparagine residue at position 8' of the HA1) in the relatively recent Victoria/3/75 strain (Figs 1, 2). The haemagglutinin RNA from another human subtype (H2)¹¹ and from an avian subtype (fowl plague)¹² are of comparable length.

The protein sequence

The haemagglutinin is synthesized as a single polypeptide chain which is subsequently modified by glycosylation and post-translational processing. The latter involves removal of an NH₂-terminal signal peptide¹⁹ and proteolytic cleavage into HA1 and HA2 to yield mature haemagglutinin and infective virus^{20,21}. Assuming that the Aichi sequence also starts with the NH₂-terminal blocking group pyroglutamic acid, as found in the Memphis H3 strain²² (a blocked amino terminus of the HA1 is a general characteristic of H3 human strains and of two putative progenitor animal strains²³⁻²⁵), the signal sequence would be 16 amino acids long, as in Victoria⁵ and Memphis^{6,22}. Within the H3 subtype, the length and hydrophobic character of the signal peptide seem to be maintained, but the exact sequence is not (Figs 1, 2). There is no variation between the positions of cysteine residues in H3 strains. Also, the sequence of the Japan and fowl plague strain can easily be aligned according to the cysteines, allowing for a minimum number of insertions/deletions between strains^{5,26}. It is therefore to be expected that the disulphide bonding pattern is the same between these three subtypes and probably in all influenza A haemagglutinins, thus giving the haemagglutinin molecules a very similar three-dimensional conformation. The disulphide bonds have recently been determined for the A/Memphis/102/72 strain²⁷.

As in the other H3 strains Victoria and Memphis and also in the H2 Japan strain, the region excised from between HA1 and HA2 consists of a single arginine residue; this situation contrasts with fowl plague haemagglutinin, where six residues are removed^{12,28}. A second hydrophobic sequence, at the NH₂ terminus of the HA2, which may have a role in interaction with and penetration of the cell membrane during infection, is highly conserved in all influenza A strains examined so far, even in a single B-type strain, and also in the NH₂ terminus of the F1 subunit of Sendai virus²⁵. Near the C terminus of HA2 there is a third hydrophobic region (amino acids 185-210 of HA2 in Fig. 1), which is involved in anchoring the haemagglutinin in the lipid viral envelope. The precise functioning of the signal interrupting the translocation of membrane proteins ('stop-transfer' sequence²⁹) remains to be established. However, in contrast to the signal sequence, this stop-transfer sequence seems to be constant within the H3 subtype in at least three different field strains.

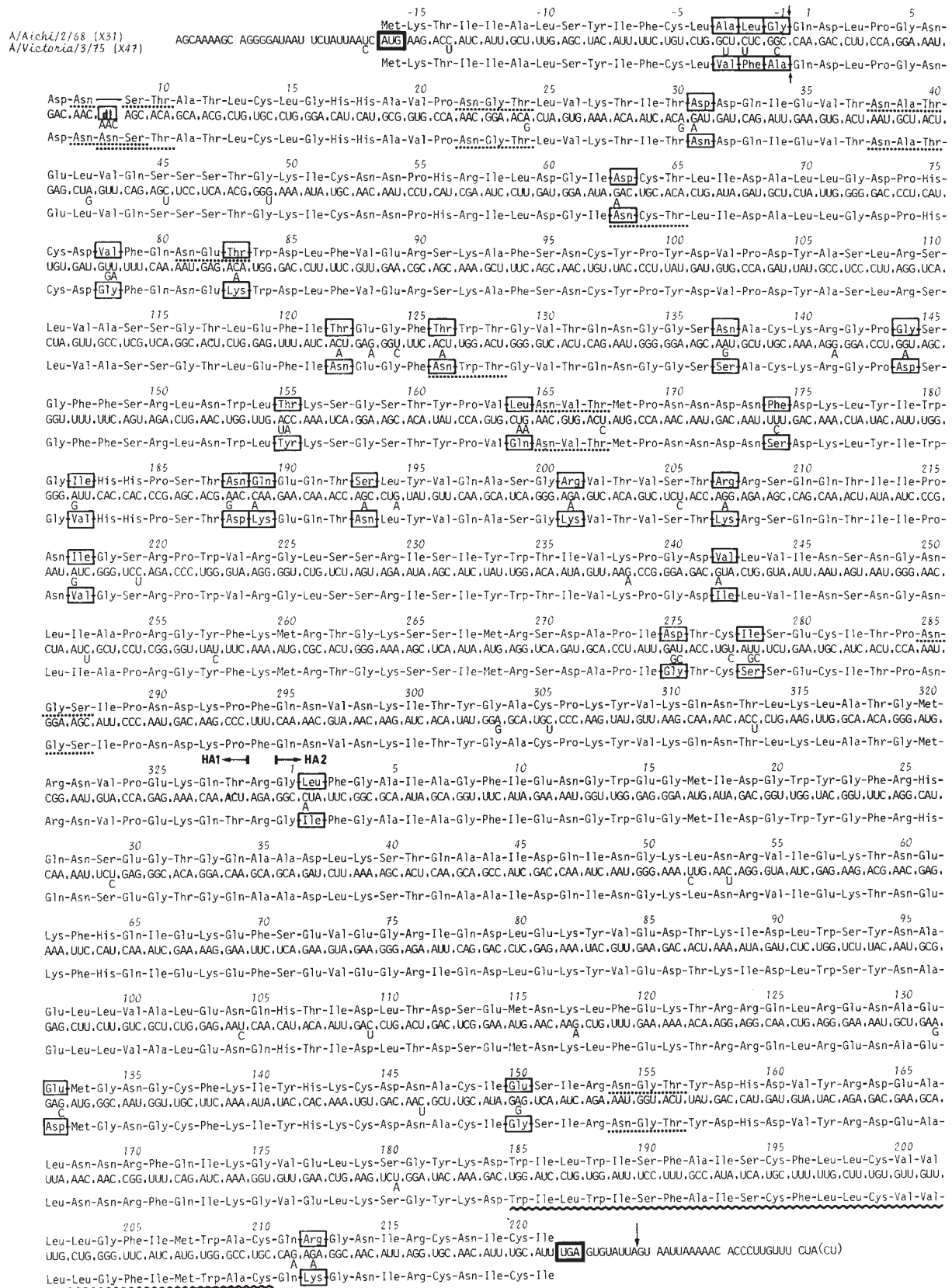


Fig. 1 Comparison of the haemagglutinin genes and deduced amino acid sequence of influenza strains A/Aichi/2/68 and A/Victoria/3/75. The nucleotide sequence of the Aichi gene is shown as the complementary viral RNA (coding strand); only the substitutions in the Victoria gene⁵ are given. The Aichi nucleotide sequence is derived from cloned DNA; however, the sequence shown from the 3' end up to the vertical arrow was missing from the cloned DNA and is assumed to be identical to the same region in Victoria (see text). The initiation and termination codons are shown in heavy boxes. Differences between amino acid sequences are shown in light boxes. Potential glycosylation sites are indicated by dotted lines. The wavy line indicates the hydrophobic region of HA2 that is embedded in the lipid membrane of the virion.

The percentage of homology in the HA1 portion is modest but usually in the same range between each two strains belonging to a different subtype (between 34 and 37%), whereas the homology is higher in HA2 (H3:H2, 50%; H3:Hav1, 66%; H2:Hav1, 52%). The fact that the HA2 is much more similar between H3 and Hav1 than between two human strains (H3 and H2) whereas the HA1 varies to approximately the same extent demonstrates how difficult it is to define the evolutionary relationship between different pandemic strains. The fraction conserved in all three subtypes (taking Aichi as the H3 prototype) is only 21% in HA1, but 42% in HA2. These values illustrate the importance of the HA1 portion to the antigenic variability of the influenza haemagglutinin.

Comparison of H3 strains: genetic drift

The nucleotide sequence of the haemagglutinin gene and the deduced amino acid sequence of the protein of the Aichi/68 and Victoria/75 strains are compared in Fig. 1. There are 67 nucleotide differences (3.8%): an insertion in the Victoria gene of three nucleotides close to the NH₂ terminus of the HA1 (coding for an extra asparagine residue) and 64 nucleotide substitutions. Of these latter changes, 63 occur in the coding region: 34 are silent nucleotide substitutions and the other 29 changes cause 28 amino acid differences (there is one case, at position 155 of the HA1, of two nucleotide changes resulting in a single amino acid change). Thus, in total there are 29 amino acid differences (including one insertion) or a 5.1% amino acid divergence accumulated over a 7-yr period (1968–75). By comparing maps of soluble tryptic peptides and amino acid composition data from nine strains isolated between 1968 and 1977 with the amino acid sequence of the A/Memphis/102/72 strain, Laver and collaborators¹⁰ found a minimum (only soluble peptides were examined) of 20 amino acid changes occurring during that period. Another conclusion was that once a residue had changed, it did not change again in the later variants, except for a few apparent reversions to the original amino acid (apparent because there is a no reason to believe that prototypes of successive epidemics as isolated from nature form a direct genealogical lineage). The same observations are made when the complete amino acid sequence of the three H3 haemagglutinin molecules, Victoria/75 (ref. 5), Memphis/72 (refs 6–8) and Aichi/68, are compared (Fig. 2). Again, a likely explanation for the apparent reversions is that, for example, the Memphis strain is not the parent of the Victoria strain. Also, at the level of silent mutation, the same phenomenon of an occasional apparent reversion is observed (not shown).

As expected, most amino acid changes were found in the HA1 part of the protein (Fig. 2, Table 1); indeed, the antigenic character is thought to reside largely, if not entirely, in this section^{30–32}. Over the 1968–75 period of antigenic drift, 22 (6.7%) of the amino acids have changed in HA1 whereas only 4 (1.8%) have been replaced in HA2. Comparing all three H3 subtypes (Fig. 2), 27 variable residues are observed in HA1 and 5 in HA2. Note that there is some clustering of the amino acid

changes in the middle third (amino acids 110–226) of the HA1 in the three strains examined. A similar differential variability between both subchains of the haemagglutinin molecule has been observed by Laver and colleagues¹⁰. Variants have been isolated which are resistant to monoclonal antiserum, and the haemagglutinins of these have probably changed at only a single amino acid position⁹. As these variants are still neutralized by total antiserum against the parent virus, it seems very likely that a more profound change (of at least two amino acids) would be required to generate a new field strain capable of overcoming the immunity of the population.

Taking into account the number of different natural variants that have been isolated and the number of epidemics reported during the period considered^{33,34}, it is reasonable to assume that a not unimportant fraction of the 22 differences in HA1 observed between Aichi/68 and Victoria/75 (probably at least one-third and possibly up to two-thirds or even more, that is, 7–15 amino acids or more) have played a part in the antigenic variability. Also, if one considers the relative abundance of differences between HA1 and HA2 to reflect the antigenic variability within HA1 above the background of tolerable mutation rates, a value of 16 residues is obtained. However, these changes may be clustered in several 'antigenic sites' (each comprising a limited number of amino acids involved in the reaction with the antibody^{9,35}) or they may influence the way the antigenic site is presented to the antibody by conformational changes. Thus, the analysis of amino acid sequences in a limited number of natural strains is not in itself sufficient for a delineation of antigenic sites.

Another approach is the analysis of variants selected with monoclonal antibodies. The Hong Kong variants studied by Laver and colleagues¹⁹ have amino acid changes at positions 54, 143, 205 and a fourth one in the region 217–224 of the HA1 (the latter two changes were observed in the variants selected with the same monoclonal antibody). There are no amino acid changes around residue 54 in the three strains compared in Fig. 2, but a change of both the adjacent residue 53 and amino acid 54 (but to a different amino acid from that in the monoclonal variant) have been described¹⁰. Again, the amino acids next to residue 143, residues 144 (Fig. 2) and 145 (ref. 10), have been found to change in field strains. Natural changes around the third variant (205) are at positions 207 (Fig. 2 and ref. 10) and 208 (ref. 10). Thus, from all cases where a direct comparison is possible, in only one instance was an amino acid change found in the same position in a monoclonal antibody-selected variant and in a field strain (residue 54), and even then the residue by which it was replaced was completely different (to a lysine and a serine, respectively), suggesting that the immunological selection in nature works differently from that in the laboratory model experiments. The reason for this (systematic) discrepancy is unknown, although it is quite clear that the immunological pressure in an infected immune individual is very different from the selection pressure working in a murine monoclonal hybridoma to select laboratory variants.

Most of the amino acid changes in these H3 haemagglutinins have been acquired by single base substitutions and are rela-

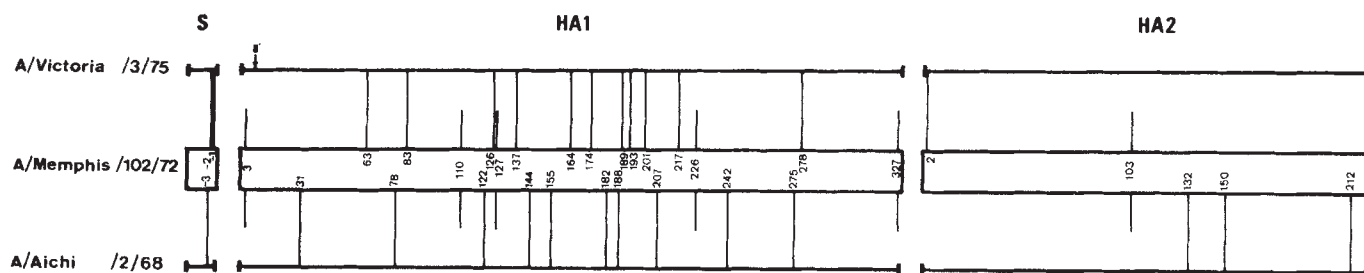


Fig. 2 Differences between the amino acid sequence of three H3 haemagglutinins derived from the virus strains A/Aichi/2/68 (the present paper), A/Memphis/102/72 (ref. 6) and A/Victoria/3/75 (ref. 5). The Memphis sequence used here is the complete one obtained by Sleight *et al.*⁸, rather than the one obtained by Ward and Dopheide^{7,8}, which lacks the hydrophobic region close to the C-terminus of the HA2. Apart from this, both sequences are essentially the same. Each difference is indicated by its number in the signal sequence (S), HA1 and HA2 portions of the molecule, respectively, and by a line connecting the strains involved. Two small lines at the same position in the Memphis strain indicate non-identity with either Aichi or Victoria (apparent reversions; see text).

Table 1 Variation observed among three H3 haemagglutinins

Comparison	Nucleotide changes leading to an amino acid change			Silent nucleotide changes		Total of nucleotide changes	
	No. of changes	% Nucleotide changes	% Amino acid changes	No. of changes	% Nucleotide changes	No. of changes	% Nucleotide changes
Aichi-Memphis	16 (4)	1.6 (0.6)	4.6 (1.8)	14 (9)	1.4 (1.4)	30 (13)	3.0 (2.0)
Memphis-Victoria	16 (2)	1.6 (0.3)	5.2 (0.9)	13 (3)	1.3 (0.4)	32 (5)	3.2 (0.8)
Aichi-Victoria	22 (4)	2.2 (0.6)	6.7 (1.8)	23 (10)	2.3 (1.5)	48 (14)	4.9 (2.1)
Total variable sites	27 (5)	2.7 (0.8)	8.2 (2.3)	24 (11)	2.4 (1.7)	54 (16)	5.5 (2.4)

The HA1 and HA2 regions are considered separately; the values for HA2 are given in parentheses. For the HA1 region the total number of nucleotides and amino acids is 984 and 328, respectively. The AAC insertion (asparagine residue 8') in the Victoria strain has been taken into account only in columns 3 (% amino acid changes) and 6 and 7 (total of nucleotide changes). The HA2 region comprises 663 nucleotides and 221 amino acids in all strains.

tively frequently observed interchanges between proteins in general³⁶. Apart from this relationship directly determined by the genetic code, chemically, the substitutions are often of a conservative nature: Gly↔Ala; Ile↔Val; Leu↔Phe; Leu↔Ile; Glu↔Asp; Arg↔Lys. There is one amino acid change (Thr→Tyr) which is both infrequently observed and of a chemically drastic nature. Whether this latter change (at position 155 of HA1) is antigenically important is not known. Not surprisingly, it is the only change caused by two nucleotide substitutions.

Glycosylation sites

All potential glycosylation sites (known to be at Asn-X-Ser or Asn-X-Thr sequences) of the Aichi haemagglutinin are identical to the sites in A/Memphis/102/72, and these glycosylation sites have recently been directly confirmed^{7,8}. Thus, it is probable that the glycosylation pattern has remained constant during the period 1968–72. However, one site (position 81 in HA1) disappears and two potential new sites (at positions 63 and 126) are created in Victoria relative to the 1968 prototype (and Memphis/72). Thus, the pattern of glycosylation in the HA1 is not invariable within the H3 subtype. The effect (if any) of this variation on the antigenic properties remains to be established. Note that the positions of the glycosylation sites are rather variable between different subtypes, except at position 154 of the HA2 chain, which has been conserved in all influenza A strains examined so far^{5–8,11,12}.

Further considerations

Not only nucleotide substitutions leading to an amino acid change but also silent mutations occur somewhat more frequently in the sequence coding for HA1 (2.3%) than in that coding for HA2 (1.5%) (Table 1). This enhanced value is statistically significant at the 5% level as tested by χ^2 analysis. If this trend is consolidated by more data, the question arises as to the meaning of this higher variability in HA1 independent of selection pressure at the protein level. It could perhaps be interpreted as base changes selected as a function of the preservation of a defined secondary structure in the viral RNA that has been more disturbed by nucleotide changes causing amino acid replacements in HA1 than HA2. A well defined secondary (and tertiary) structure may be important in the formation of the nucleoprotein complex or for packaging the eight RNAs in a molar ratio in the particle. Another case where the importance of secondary structure in the genome of RNA viruses has been well documented involves the RNA phages, where it has been shown that the number of base changes between RNA phages is 2.5 times lower in double-stranded regions than in single-stranded regions³⁷.

If one considers the actual number of silent and non-silent single base changes in the 61 codons for the 20 amino acids in the HA1 and the HA2 coding sequence and compares these values with those that would be expected assuming that single base replacements are acquired on a random basis, the differences are quite striking. According to this assumption, 25.5% of the substitutions would be expected to be neutral and 74.5% would cause an amino acid change³⁸. In reality, one finds in the HA1: silent changes, 23; non-silent changes, 20; non-silent changes expected taking the silent ones as the 25.5% value, 67. In the HA2 region these values are: silent changes, 10; non-silent changes, 4; non-silent changes expected, 29. This means that in HA1, 70% of the potentially non-silent mutations would be selected against; this percentage is even higher (86%) in HA2. A similar comparison between RNA phages leads to a value of 93% (ref. 37). All these deviations from randomness are statistically meaningful at the 0.1% level (χ^2 test). Although these estimates are not very accurate (it is not valid to assume an equal probability for a nucleotide to change in any direction), they are, nevertheless, a relative measure of the selection pressure to maintain the amino acid sequence in a molecule or its sub-regions.

Considering the nature of the nucleotide substitutions, transitions are much more frequent than transversions: of the 64 single-base substitutions between Aichi and Victoria, 49 are transitions and only 15 are transversions. Similarly, comparison of nucleotide sequences from two RNA bacteriophages, MS2 and R17, reveals 29 transitions and 5 transversions³⁷. Even in the (at least at first approximation) neutral third-letter position of fourfold degenerate codons, 12 out of 15 changes are transitions between the influenza strains, and all 11 changes are transitions in the comparison between the two phages. It is quite possible that these transitions arise during replication by the formation of G·U instead of the normal G·C or A·U interactions³⁷.

Finally, it is too early to understand the basis of drift from the sequence information available for the haemagglutinins of natural H3N2 strains and of variants selected by monoclonal antibodies. The three-dimensional structure of the X31 haemagglutinin will probably soon be solved³⁹, especially as the primary amino acid sequence is now available for reference. This may then, in combination with the available and forthcoming sequence information, provide a solid framework for understanding the variability of the virus within a subtype, the mechanism of antigenic drift and the number and distribution of antigenic sites on the haemagglutinin molecule.

We thank Dr J. Skehel for a gift of X31 RNA and W. Kuziel for help with the manuscript. This research was supported by grants from the Fonds voor Kollektief Fundamenteel Onderzoek and the Gekoncerteerde Akties of the Belgian Ministry of Science. D. H. thanks the IWONL and M.V. thanks the NFWO for fellowships.

Received 16 April; accepted 17 June 1980.

1. Laver, W. G. & Webster, R. G. *Br. med. Bull.* **35**, 29–33 (1979).
2. Palese, P. & Schulman, J. L. *Proc. natn. Acad. Sci. U.S.A.* **73**, 2142–2146 (1976).
3. Scholtissek, C., Harms, E., Rohde, W., Orlich, M. & Rott, R. *Virology* **74**, 332–344 (1976).
4. McGeoch, D. J., Fellner, P. & Newton, C. *Proc. natn. Acad. Sci. U.S.A.* **73**, 3045–3049 (1976).

5. Min Jou, W. *et al. Cell* **19**, 683–696 (1980).

6. Sleight, M. J., Both, G. W., Brownlee, G. G., Bender, V. J. & Moss, B. A. in *Structure and Variation in Influenza Virus* (eds Laver, W. G. & Air, G. M.) 69–79 (Elsevier, Amsterdam, 1980).

7. Ward, C. W. & Doppeide, T. A. in *Structure and Variation in Influenza Virus* (eds Laver, W. G. & Air, G. M.) 27–38 (Elsevier, Amsterdam, 1980).

8. Doppeide, T. A. & Ward, C. W. in *Structure and Variation in Influenza Virus* (eds Laver, W. G. & Air, G. M.) 21–26 (Elsevier, Amsterdam, 1980).

9. Laver, W. G. *et al.* *Virology* **98**, 226–237 (1979).
10. Laver, W. G., Air, G. M., Doppeide, T. A. & Ward, C. W. *Nature* **283**, 454–457 (1980).
11. Gething, M.-J., Bye, J., Skehel, J. & Waterfield, M. in *Structure and Variation in Influenza Virus* (eds Laver, W. G. & Air, G. M.) 1–10 (Elsevier, Amsterdam, 1980).
12. Porter, A. G. *et al.* *Nature* **282**, 471–477 (1979).
13. Devos, R., Gillis, E. & Fiers, W. *Eur. J. Biochem.* **62**, 401–410 (1976).
14. Devos, R., van Emmelo, J., Contreras, R. & Fiers, W. *J. molec. Biol.* **128**, 595–619 (1979).
15. Grunstein, M. & Hogness, D. S. *Proc. natn. Acad. Sci. U.S.A.* **72**, 3961–3965 (1975).
16. Smith, H. O. & Birnstiel, M. L. *Nucleic Acids Res.* **3**, 2387–2398 (1976).
17. Maxam, A. M. & Gilbert, W. *Proc. natn. Acad. Sci. U.S.A.* **74**, 560–564 (1977).
18. Efstratiadis, A., Kafatos, F. C. & Maniatis, T. *Cell* **10**, 571–585 (1977).
19. McCauley, J. *et al.* *FEBS Lett.* **108**, 422–426 (1979).
20. Lazarowitz, S. G. & Chopin, P. W. *Virology* **68**, 440–454 (1975).
21. Klenk, H.-D., Rott, R., Orlich, M. & Blödorn, J. *Virology* **68**, 426–439 (1975).
22. Ward, C. W. & Doppeide, T. A. *Virology* **103**, 37–53 (1980).
23. Laver, W. G. & Webster, R. G. *Virology* **81**, 482–485 (1977).
24. Ward, C. W. & Doppeide, T. A. *Br. med. Bull.* **35**, 51–56 (1979).
25. Waterfield, M. D., Espelie, K., Elder, K. & Skehel, J. J. *Br. med. Bull.* **35**, 57–63 (1979).
26. Brownlee, G. G. in *Structure and Variation in Influenza Virus* (eds Laver, W. G. & Air, G. M.) 385–390 (Elsevier, Amsterdam, 1980).
27. Doppeide, T. A. & Ward, C. W. *FEBS Lett.* **110**, 181–183 (1980).
28. Klenk, H.-D. in *Structure and Variation in Influenza Virus* (eds Laver, W. G. & Air, G. M.) 213–222 (Elsevier, Amsterdam, 1980).
29. Blobel, G. *et al.* *Soc. exp. Biol. Symp.* **33**, 9–36 (1979).
30. Jackson, D. C., Doppeide, T. A., Russell, R. J., White, D. O. & Ward, C. W. *Virology* **93**, 458–465 (1979).
31. Jackson, D. C., Russell, R. J., Ward, C. W. & Doppeide, T. A. *Virology* **89**, 199–205 (1978).
32. Potter, C. W. & Oxford, J. S. *Br. med. Bull.* **35**, 69–75 (1979).
33. Pereira, M. S. in *Influenza Virus, Vaccines and Strategy* (ed. Selby, P.) 25–31 (Academic, London, 1976).
34. Pereira, M. S. *Br. med. Bull.* **35**, 9–14 (1979).
35. Atassi, M. Z. & Smith, J. A. *Immunochimistry* **15**, 609–610 (1978).
36. Dayhoff, M. O., Schwartz, R. M. & Orcutt, B. C. in *Atlas of Protein Sequence and Structure* Vol. 5, Suppl. 3 (ed. Dayhoff, M. O.) 345–352 (The National Biomedical Research Foundation, Silver Springs, Maryland, 1978).
37. Min Jou, W. & Fiers, W. *J. molec. Biol.* **106**, 1047–1060 (1976).
38. Jukes, T. H. & King, J. L. *Nature* **281**, 605–606 (1979).
39. Wilson, I. A., Skehel, J. J. & Wiley, D. C. in *Structure and Variation in Influenza Virus* (eds Laver, W. G. & Air, G. M.) 339–348 (Elsevier, Amsterdam, 1980).

RNA splicing generates a variant light chain from an aberrantly rearranged κ gene

Edmund Choi*, Michael Kuehl† & Randolph Wall*

* Molecular Biology Institute and Department of Microbiology and Immunology, UCLA School of Medicine, Los Angeles, California 90024
† Department of Microbiology, University of Virginia School of Medicine, Charlottesville, Virginia 22908

Both C_{κ} regions in MPC 11 cells are rearranged into active transcription units, one producing a normal κ chain and the other an internally deleted κ fragment lacking a V region. The gene coding for the κ fragment mRNA is aberrantly rearranged and lacks a site for $V \rightarrow C_{\kappa}$ splicing. An alternative splicing event which deletes the V region from the nuclear RNA precursor generates the κ fragment mRNA.

IMMUNOGLOBULIN light-chain variable (V) and constant (C) regions are rearranged in the formation of active genes (that is, transcription units)^{1–4}. These somatic rearrangements join a specific germ-line V region with a J region while retaining the intervening sequence (intron) separating the J and C regions in germ-line DNA^{1–4}. Some myeloma cells also contain C regions arranged differently from either the germ-line pattern or the pattern in the active transcription unit producing immunoglobulin light chains^{5,6}. The detailed structure of such rearranged C regions has not been determined. These presumably aberrant DNA arrangements have been proposed to be the mechanism for allelic exclusion in immunoglobulin-producing cells. However, it has not been shown whether or not such aberrantly rearranged C regions are transcribed. Indeed, the relationship of DNA rearrangements to the activation of immunoglobulin genes for transcription remains to be resolved.

Immunoglobulin messenger RNAs, like virtually every eukaryotic mRNA now studied, are processed from larger nuclear RNA precursors by RNA splicing^{7–10}. It seems that RNA splicing may be an obligatory step for the generation of stable mRNA^{11–13}. One clue to the signals possibly used in RNA splicing has emerged from the nucleotide sequencing of cloned eukaryotic genes containing introns. All eukaryotic genes now studied (including several immunoglobulin light- and heavy-chain genes, with one notable exception) have the characteristic nucleotides 5′-/GT-intron-AG/-3′ at the junctures between introns and exons (summarized in ref. 14). Compilation of subsequent sequences has shown that the 5′ (upstream) and 3′ (downstream) ends of introns approximate to the following consensus or 'optimal' splice sites^{14–16}

5′-exon-AG/GTAAGTA—intron—
—TTTTYTTTTTCTTNCAG/G-exon-3′

These upstream and downstream splice sites flanking an intron presumably must be brought together for RNA splicing to occur. One of the abundant, small nuclear RNA (snRNA) species called U-1 (ref. 17) is exactly complementary to these 'optimal' sequences at RNA splicing sites. A mechanism for RNA splicing has been proposed in which U-1 snRNA is the recognition

component of the nuclear RNA splicing enzyme complex and hybridizes with both ends of an intron so as to align them in exact register for RNA splicing^{15,16}. However, the significance of these presumptive signals and the molecular mechanisms of RNA splicing remain to be elucidated.

We have approached these questions by investigating how cloned MPC 11 myeloma cells simultaneously produce both a normal κ light-chain containing precursor, P, V, J and C_{κ} regions, and a variant κ light-chain fragment containing P and C_{κ} regions but lacking V and J regions^{18,19}. Combining nuclear RNA and DNA mapping approaches, we find that the two κ mRNAs in MPC 11 cells are independently transcribed from two separate transcription units with rearranged C_{κ} regions. Aberrant rearrangement of one C_{κ} region in MPC 11 cells generates the transcription unit for the κ fragment light chain. The nuclear RNA transcript of this aberrantly rearranged C_{κ} region undergoes an unusual RNA splicing step which generates the κ fragment mRNA. The detailed structure of this variant MPC 11 C_{κ} gene is presented in the accompanying article²⁰.

MPC 11 cells contain two rearranged C_{κ} regions

The expression of two different κ light chains in MPC 11 cells suggests that there must be two differently rearranged C_{κ} regions. This prediction was confirmed by Southern mapping²¹ of the C_{κ} regions in *Bam*HI-cleaved DNA from MPC 11 (clone 45.6) and spleen cells. Hybridization with a cloned cDNA probe (pL21-5, 22) which contains the C_{κ} region reveals that 45.6 cells contain strongly hybridizing fragments of 3.0 and 7.7 kilobases (Fig. 1). Spleen DNA which has non-rearranged C_{κ} regions contains a single C_{κ} band at 14 kilobases. Weakly hybridizing fragments of 5.1, 5.5 and 5.7 kilobases, present in all the cell lines and spleen DNA, are probably due to a small amount (≈ 50 nucleotides) of MOPC 21 V-region sequences in the pL21-5 probe²² which react with MOPC 21 and related V regions in the DNA. The MOPC 21 V-region does not cross-hybridize with either the MPC 11 or MOPC 321 V regions (E.C., unpublished results).

To determine which C_{κ} restriction fragments correspond to the normal and fragment MPC 11 light chains, we examined a variant (clone NP.2) which expresses only the κ fragment but not the normal light chain. In contrast to those of 45.6 cells, the *Bam*HI digests of NP.2 cells reveal only a single strongly hybridizing fragment of 3.0 kilobases. As this is the only C_{κ} fragment in NP.2 cells, it must represent the rearranged C_{κ} region which encodes the κ fragment chain. Apparently, the 7.7 *Bam*HI fragment represents the C_{κ} region which encodes the normal κ light chain in MPC 11 cells, because the loss of normal light-chain expression in NP.2 cells is correlated with the loss of this fragment. Thus, the loss of normal light-chain expression in the NP.2 variant clone, which was isolated from MPC 11 cells by nitrosoguanidine mutagenesis, seems to be due to the deletion of the rearranged C_{κ} region coding for this chain.

Nuclear RNA species derived from the normal and fragment genes

Previous results from our laboratory²³ and from Schibler *et al.*⁹ revealed multiple κ light-chain nuclear RNA species in MPC 11 cells. To distinguish between the nuclear RNA processing pathways for normal κ and the κ fragment mRNAs, we compared the κ nuclear RNA species in wild-type MPC 11 45.6 cells making both κ chains with those in the NP.2 variant line which synthesized only the C_{κ} fragment chains.

Polyadenylated nuclear RNA was fractionated on methyl mercury hydroxide gels, transferred onto diazophenylthioether (DPT) paper ('RNA blots', ref. 24) and hybridized with nick-translated C_{κ} -gene probe (Fig. 2, panel *a*). The 45.6 cells contain five distinct C_{κ} nuclear RNA species: 6.5, 5.6, 3.6, 1.2 and 0.8 kilobases. The 1.2- and 0.8-kilobase species represent the normal κ light-chain and κ fragment mRNAs, respectively^{9,23}. These κ nuclear RNA species are in approximate agreement with those previously reported^{9,23} except that the largest band is here resolved into 6.5- and 5.6-kilobase species. When the C_{κ} -gene probe was hybridized to nuclear RNA from NP.2 cells, only the 3.6- and 0.8-kilobase species were seen (Fig. 2, panel *a*).

These findings outline the two different κ light-chain RNA processing pathways in MPC 11 cells. No larger κ nuclear RNA species are detected either by RNA blots or in brief pulse-labelling experiments (E. C., M. K. and R. W., unpublished results). Although it is not formally proven here, we presume that the polyadenylated 6.5- and 3.6-kilobase species represent the primary transcripts for the κ light-chain and κ fragment mRNAs, respectively. The definition of the sequences in these presumptive primary transcripts provides a means of further defining the structures of the two active κ transcription units in MPC 11 cells.

To characterize these κ transcripts further, we isolated two specific restriction fragments of the C_{κ} -gene clone^{2,3} (see Fig. 2, inset). One contains a portion of the intron between the J regions and C_{κ} region (designated IS), the other fragment encompasses the entire J-region cluster (designated J; see Fig. 2, inset). The intron probe hybridized to the 6.5- and 5.6-kilobase precursors in 45.6 cells (Fig. 2, panel *b*). It also hybridized to the 3.6-kilobase species in both 45.6 and NP.2 cells, indicating that this nuclear RNA species contains sequences from the J $\rightarrow$ C_{κ} intron in normal κ light-chain transcription units. As expected, this probe did not hybridize with either the 1.2- or the 0.8-kilobase mRNA species.

We next determined which nuclear RNA species contained J-region sequences. As expected, the restriction fragment containing all five known κ J regions hybridized with the 6.5-, 5.6- and 1.2-kilobase κ RNAs in 45.6 cells (Fig. 2, panel *c*). In contrast, the 3.6- and 0.8-kilobase κ species in 45.6 or NP.2 cells did not hybridize with the J probe (Fig. 2, panel *c*), suggesting that the transcription unit for the κ fragment does not contain J regions.

It has previously been shown that the MPC 11 κ fragment mRNA contains a MOPC 321-like hydrophobic precursor (P) region linked to the C_{κ} region¹⁹. To determine if the nuclear

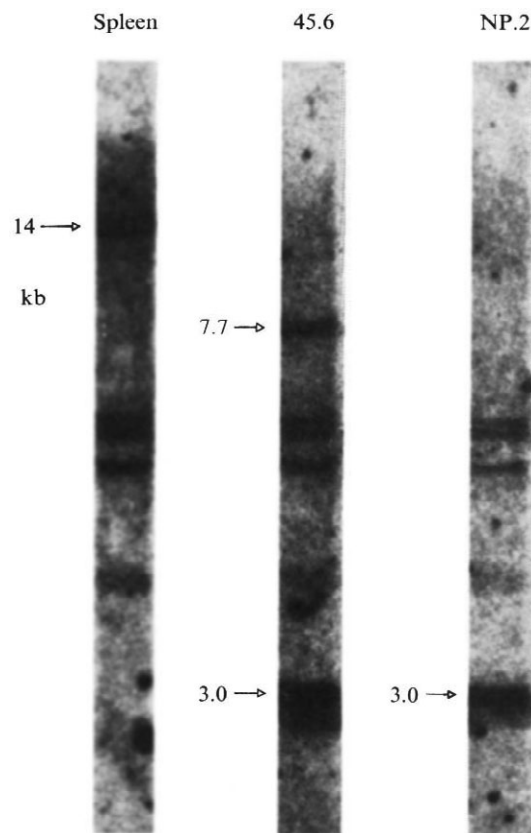
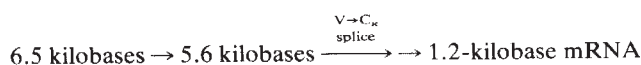


Fig. 1 Southern hybridization analysis of the κ constant region specific *Bam*HI DNA fragments of spleen and MPC 11 wild-type and variant plasmacytomas. Fragments generated by *Bam*HI digestion of BALB/c spleen, MPC 11 clone 45.6, and MPC 11 clone NP.2 DNAs were fractionated by electrophoresis in 0.75% gels, alkali denatured, neutralized and transferred to nitrocellulose filters²¹. The filters were hybridized to 0.2 μ g of nick-translated ³²P-labelled pL21-5 plasmid DNA (1.5×10^8 c.p.m. per μ g) and subjected to autoradiography. The pL21-5 cDNA plasmid probe contains approximately 550 nucleotides of MOPC 21 κ mRNA sequence, including all the 3'-untranslated, constant region and J-region sequences as well as 50 nucleotides of MOPC 21 V-region sequences²². Fragment lengths were determined by comparison with ethidium bromide-stained *Eco*RI and *Hind*III fragments of phage λ phage DNA. Lane *a*, spleen DNA; lane *b*, MPC 11 clone 45.6; lane *c*, MPC 11 clone NP.2 DNA.

RNA precursor to the C_{κ} fragment mRNA also contains the MOPC 321 V region, we hybridized a nick-translated restriction fragment containing the V region of MOPC 321 to nuclear RNA from 45.6 and NP.2 cells. Only the 3.6-kilobase species in both cell lines contains a V region which hybridizes with the MOPC 321 variable region (Fig. 2, panel *d*). This probe did not hybridize to the 6.5- or 5.6-kilobase RNA precursors in 45.6 cells, establishing that there are no MOPC 321 V-region sequences on the precursors to the normal κ light chain of MPC 11 cells.

The two independent κ nuclear RNA processing pathways are clearly defined by these results. Those RNA precursors in the normal κ light-chain processing pathway are identified by the presence of J-region sequences, and those in the κ fragment pathway are distinguished by the presence of MOPC 321 V-region sequences.

The processing pathway for the full-length κ mRNA is



In the transcription unit for the normal MPC 11 κ light chain, the V region is joined to J₄ and contains a 3.5-kilobase intron separating the joined V + J from the C_{κ} region^{3,4}. This intron is removed in the processing step from the 5.6-kilobase nuclear RNA to the 1.2-kilobase mRNA, apparently with another

Fig. 2 Analysis of nuclear RNA species in 45.6 and NP.2 cells. Nuclear RNA was prepared from detergent-washed nuclei³¹. Poly(A)⁺ nuclear RNA was fractionated on 1% agarose gels containing 5.0 mM methyl mercury hydroxide³². Nuclear RNA was loaded at 20 µg per lane. All operations involving methyl mercury hydroxide were carried out using appropriate safety measures. The RNA transfer was carried out as described by Wahl *et al.*³³ except that DPT paper was used instead of diazobenzyloxymethyl paper. Following transfer, the DPT paper was hybridized with nick-translated probes and washed in conditions described by Wahl *et al.*³³. The restriction fragments used for nick translation were isolated from agarose gels²³. The specific activities of the ³²P-labelled probes ranged from 7×10^7 to 2×10^8 c.p.m. per µg; $2-3 \times 10^6$ c.p.m. were hybridized to each lane. The following probes were each hybridized to a pair of lanes containing the nuclear RNA from 45.6 and NP.2 cells. Panel a, C_κ embryo genome clone. The 6.5-kilobase *Xba*I-*Hind*III restriction fragment³ from the C_κ clone is outlined in the inset below the RNA lanes. Panel b, intervening sequence probe. A 1.1-kilobase *Xba*I-*Hind*III restriction fragment designated IS in the C_κ embryo gene was used. Panel c, J-region probe. A 1.7-kilobase *Hind*III/*Xba*I restriction fragment designated J in the diagram was used. Panel d, 321 variable region probe. An *Ava*I restriction fragment containing the 321 variable region was isolated from a recombinant cDNA clone. Filters in panels b, c and d rehybridized with the C_κ embryonic probe used in lane a showed that all the κ nuclear species were present on these filters (data not shown). The following mouse rRNAs were used to calculate the sizes of hybridized nuclear RNA bands 45s (14.1 kilobases), 32s (6.6 kilobases), 28s (5.2 kilobases) and 18s (2.1 kilobases)³⁴.

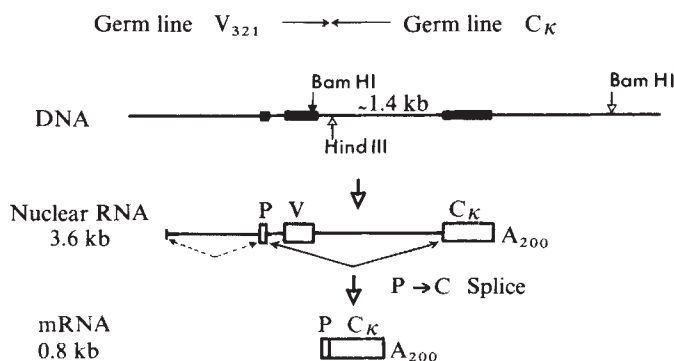
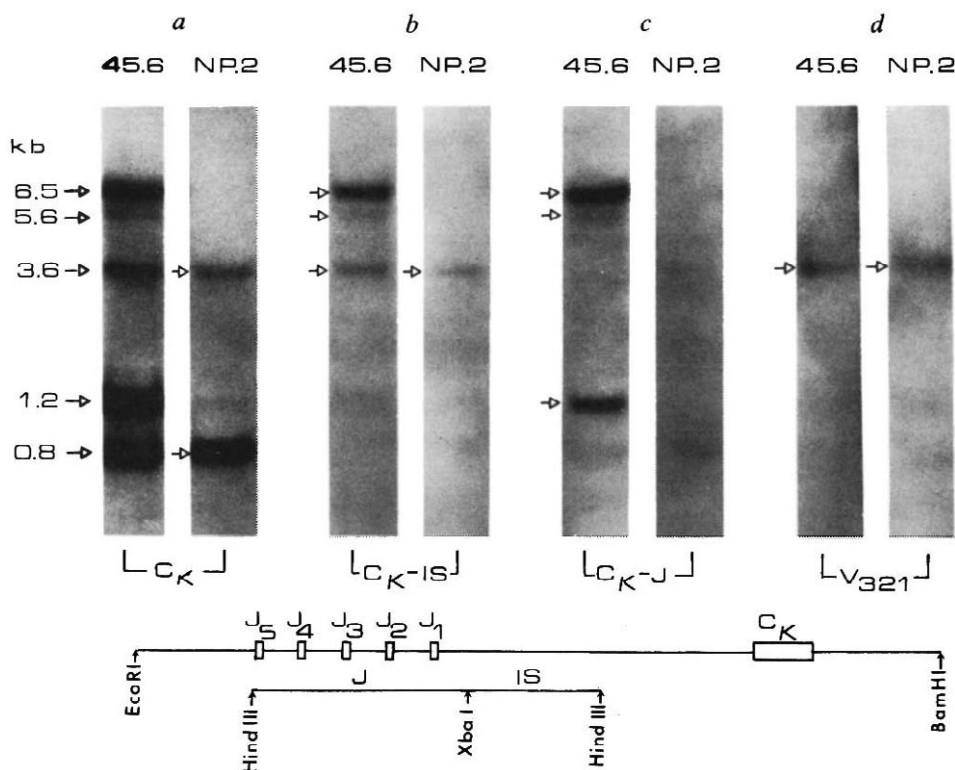
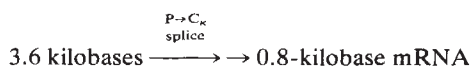


Fig. 3 The transcription unit and nuclear RNA processing pathway for the MPC 11 C_κ fragment mRNA. Polyadenylation of the 3.6-kilobase nuclear RNA precedes RNA splicing. Other details of the model are discussed in the text.

undefined intron amounting to approximately 1 kilobase. This V → C_κ splicing event may proceed through a rapidly labelled intermediate of approximately 2.1 kilobases which is readily detected in pulse-labelling experiments²³ but not in RNA blots, as the latter technique primarily detects stable species of nuclear RNA.

The processing pathway for the κ fragment mRNA seems to be



Because the 3.6-kilobase nuclear RNA precursor lacks a J region and the germ-line MOPC 321 V region lacks an upstream splice site, this V region cannot be spliced to the C_κ region. The κ fragment chain lacks the last three carboxy-terminal amino acids of the MOPC 321 P region^{18,19}. This suggests that the normal splicing site at the 3' end of the P region is spliced to the site at

the 5' side of the C_κ region, thereby generating the fragment light-chain mRNA (Fig. 3). It is not known if the P region is alternatively spliced to the 5' end of the 321 V-region gene. The RNA product of such a splice, removing only approximately 0.2 kilobases, would probably not be distinguished in RNA blots. Moreover, the product RNA might be rapidly degraded in the nucleus. The efficiency of P → C_κ splicing is not known. However, it is likely that this unusual splice occurs at a comparatively high frequency, because the 3.6-kilobase RNA species and the κ-fragment mRNA comprise, respectively, some 20% of the κ RNA in the nucleus and cytoplasm of 45.6 cells. The P → C_κ splice removes approximately 2 kilobases. An additional splice removing approximately 0.8 kilobases also apparently occurs in the generation of the κ fragment mRNA. This splice presumably involves 5'-untranslated leader sequences (Fig. 3).

Structure of the κ fragment transcription unit

The fragment transcription unit apparently contains P, V and C regions, but lacks J regions. Moreover, it contains the intervening sequence up to 1.1 kilobases from the 5' side of the C_κ region (that is, sequences homologous to the *Xba*I/*Hind*III fragment; Fig. 2, inset). A model consistent with these results would predict that the embryonic MOPC 321 V region is joined to a point in the intervening sequence preceding the embryonic C_κ gene (including the *Hind*III site 1.1 kilobases from the C_κ region), such that the J regions are deleted. The presence of the fragment κ gene on a 3.0-kilobase *Bam*HI restriction fragment allows us to localize the rearrangement further. A *Bam*HI restriction site is located 1.2 kilobases 3' to the embryonic C_κ gene^{2,3}. The sequences flanking the 3' end of the C_κ gene are conserved during the somatic rearrangement of the DNA to form the active immunoglobulin gene^{2,3,6}. Assuming that the C_κ gene amounts to 0.5 kilobases (including the contiguous 3'-untranslated region), the *Bam*HI site 5' to the C_κ gene is approximately 1.4 kilobases away. No *Bam*HI site exists in this region in the embryonic C_κ gene^{2,3}. Therefore, this *Bam*HI site

must have originated on the 3' end of the rearranged MOPC 321 V region². A likely *Bam*HI site is found at amino acid 96 in the germ-line MOPC 321 variable region². This would place the rearranged 321 V region approximately 1.4 kilobases away from the C_κ region (Fig. 3). This predicted arrangement of the MPC 11 κ fragment transcription unit has been confirmed by direct nucleotide sequence by Seidman and Leder²⁰.

Discussion

These results clearly show that both C_κ alleles can be rearranged into concomitantly active transcription units. However, with regard to allelic exclusion, it should be noted that these MPC 11 cells produce only a single functional secreted κ chain. Because the aberrant rearrangement which generates the κ fragment occurs by recombination between homologous germ-line DNA sequences²⁰, it is likely that other immunoglobulin-producing cells will also express V regions joined at this intervening sequence site. Certainly, the demonstration that such rearranged C_κ genes are transcribed suggests that the activation of immunoglobulin C_κ genes for transcription does not necessarily result from the normal V→J joining event. Recent findings indicate that in certain myeloma cell lines, an unrearranged C_κ gene is apparently transcribed concomitantly with the rearranged C_κ gene which generates mRNA²⁵. This suggests that DNA rearrangement may not initiate the transcription of an immunoglobulin gene but may rather generate a complete transcription unit bounded by an initiation site for transcription (contributed by the V-region segment) and a poly(A) addition site (contributed by the C-region segment). These two elements may be sufficient to delineate a transcription unit capable of

producing a stable, polyadenylated primary transcript which can be processed to immunoglobulin mRNA.

The MPC 11 κ fragment mRNA contains two exons (P, C_κ) which are not normally spliced together in immunoglobulin mRNA. Such shuffling of exons presumably does not occur very frequently in normal mRNA splicing. However, RNA splicing can apparently bypass a mutant splicing site provided that the RNA precursor contains other introns with functional splicing sites. In such precursors, as in the MPC 11 κ fragment, such an event could result in the deletion of a complete exon.

This seems to be the most likely explanation for the large number of immunoglobulin heavy-chain deletion variants which have lost internal domains or the hinge region^{26,27}. A particularly striking example of such a heavy chain deletion variant is the IF2, γ₁ chain, which lacks the C_H1 domain and has the V_H region joined directly to the hinge²⁸. A γ₁ heavy-chain gene which may resemble the IF2 variant gene has been cloned from MOPC 21 myeloma tumour DNA²⁹. This MOPC 21 γ₁ clone is identical in nucleotide sequence to another γ₁ gene clone from embryo DNA³⁰ except that the trinucleotide, GTG, which directly follows the C_H1 domain, is deleted^{29,30}. Presumably, RNA transcripts of this MOPC 21 γ₁ gene lack a functional splice site at the C_H1/intron juncture and would be expected to splice the V_H region to the hinge, producing an internally deleted variant chain like IF2. Future studies on immunoglobulin deletion variants should serve to define further the signals and steps in RNA splicing.

These studies were supported by NIH research grants AI13410, CA12900 and AI12525. E. C. was supported by NIH Training Grant CA 09056. DPT paper was developed by Brian Seed, C_κ embryo clone was made available to us by P. Leder, and the MOPC 321 cDNA clone was from S. Tonegawa.

Received 4 April; accepted 18 June 1980.

- Brack, C., Hiram, M., Lenhard-Schuller, R. & Tonegawa, S. *Cell* **15**, 1-14 (1978).
- Sakano, H., Hüppi, K., Heinrich, G. & Tonegawa, S. *Nature* **280**, 288-294 (1979).
- Max, E. E., Seidman, J. G. & Leder, P. *Proc. natn. Acad. Sci. U.S.A.* **76**, 3450-3454 (1979).
- Seidman, J. G., Max, E. E. & Leder, P. *Nature* **280**, 370-375 (1979).
- Lenhard-Schuller, R., Hohn, B., Brack, C., Hiram, M. & Tonegawa, S. *Proc. natn. Acad. Sci. U.S.A.* **75**, 4709-4713 (1978).
- Seidman, J. G. & Leder, P. *Nature* **276**, 790-795 (1978).
- Gilmore-Hebert, M. & Wall, R. *Proc. natn. Acad. Sci. U.S.A.* **75**, 342-345 (1978).
- Gilmore-Hebert, M. & Wall, R. *J. molec. Biol.* **135**, 879-891 (1979).
- Schibler, U., Marcu, K. B. & Perry, R. P. *Cell* **15**, 1495-1509 (1978).
- Perry, R. P., Kelley, D. E. & Schibler, U. *Proc. natn. Acad. Sci. U.S.A.* **76**, 3678-3682 (1979).
- Gruss, R., Lai, C.-J., Dahr, R. & Khoury, G. *Proc. natn. Acad. Sci. U.S.A.* **76**, 4317-4321 (1979).
- Hamer, D. H. & Leder, P. *Cell* **18**, 1299-1302 (1979).
- Berg, P. in *Mobilization and Reassembly of Genetic Information* (eds Scott, W. A., Schultz, J., Werner, R. & Joseph, D. R.) (Academic, New York, in the press).
- Seif, I., Khoury, G. & Dhar, R. *Nucleic Acids Res.* **6**, 3387-3398 (1979).
- Rogers, J. H. & Wall, R. *Proc. natn. Acad. Sci. U.S.A.* **77**, 1877-1879 (1980).
- Lerner, M. R., Boyle, J. A., Mount, S. M., Wolin, S. L. & Steitz, J. A. *Nature* **283**, 220-224 (1980).
- Ro-Choi, T. S. & Henning, D. *J. biol. Chem.* **252**, 3814-3820 (1979).
- Kuehl, W. M. *et al. Cell* **5**, 139-147 (1975).
- Rose, S. M., Kuehl, W. M. & Smith, G. P. *Cell* **12**, 453-462 (1977).
- Seidman, J. G. & Leder, P. *Nature* **286**, 779-783 (1980).
- Southern, E. M. *J. molec. Biol.* **98**, 503-517 (1975).
- Wall, R. *et al. Nucleic Acids Res.* **5**, 3113-3128 (1978).
- Komaromy, M., Clarke, P., Choi, E., Kuehl, M. & Wall, R. in *Eukaryotic Gene Regulation* Vol. 14 (eds Axel, R., Maniatis, T. & Fox, C. F.) 421-431 (Academic, New York, 1979).
- Alwine, J. C., Kemp, D. J. & Stark, G. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5350-5354 (1977).
- Perry, R. P. *et al., Proc. natn. Acad. Sci. U.S.A.* **77**, 1937-1941 (1980).
- Franklin, E. C. & Frangione, B. in *Contemporary Topics in Molecular Immunology* (ed. Inman, F. P.) 89-126 (Plenum, New York, 1975).
- Franklin, E. C., Prelli, F. & Frangione, B. *Proc. natn. Acad. Sci. U.S.A.* **76**, 452-456 (1978).
- Adetugbo, K. *J. biol. Chem.* **253**, 6076-6080 (1978).
- Sakano, H. *et al. Nature* **277**, 627-633 (1979).
- Honjo, T. *et al. Cell* **18**, 559-568 (1979).
- Fedoroff, N., Wellauer, P. K. & Wall, R. *Cell* **10**, 597-610 (1977).
- Bailey, J. M. & Davidson, N. *Analyt. Biochem.* **70**, 75-85 (1976).
- Wahl, G. M., Stern, M. & Stark, G. R. *Proc. natn. Acad. Sci. U.S.A.* **76**, 3683-3687 (1979).
- Wellauer, P. K., Dawid, I. B., Kelley, D. E. & Perry, R. P. *J. molec. Biol.* **89**, 397-407 (1974).

A mutant immunoglobulin light chain is formed by aberrant DNA- and RNA-splicing events

J. G. Seidman & Philip Leder

Laboratory of Molecular Genetics, National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, Maryland 20205

A mutant immunoglobulin gene has been formed by an abnormal (non V/J) recombination event such that abnormal RNA splicing is required to form a mutant light chain. The structure of the gene suggests that the small palindrome thought to be involved in V/J joining also provides the basis for this abnormal DNA recombination and that the absence of a J segment and RNA splice signal allows an abnormal RNA splicing reaction to occur.

THE generation of antibody diversity depends on a specialized form of genetic recombination that operates in somatic cells during lymphocyte differentiation¹⁻⁴. The combinational power of joining one of several hundred variable region genes to one of

the four J-region segments encoded close to the constant-region gene creates literally thousands of unique genetic possibilities⁵⁻⁹. Recent DNA cloning and nucleotide sequences analyses have revealed many of the molecular details of this

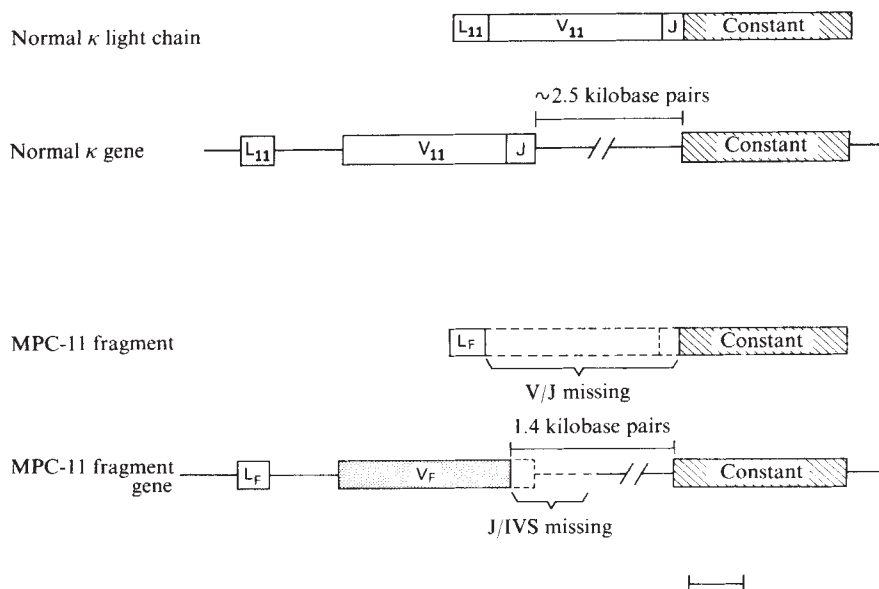


Fig. 1 The structure of the MPC-11 light chains and their genes. The MPC-11 cell synthesizes two κ light chains—a normal κ light chain and the MPC-11 fragment. We presume that the normal κ light chain is encoded by a normal κ gene, while the MPC-11 fragment is encoded by the MPC-11 fragment gene. The leader of the MPC-11 normal κ light chain (L_{11}) and the leader of the MPC-11 fragment (L_F) do not have the same amino acid sequence^{12,13}. The amino acid sequence of L_F is identical to the leader of the MOPC-321 light chain^{12,13}. Scale bar, 0.1 kilobase pairs.

process, including a plausible molecular basis for the special recombination event^{8,9}. They also provide an accurate picture of an active κ immunoglobulin gene⁷. In its final form, the κ light-chain gene consists of three separate coding elements separated by two intervening sequences of DNA. Obviously, the formation of a light chain requires both DNA- and RNA-splicing reactions.

Kuehl and Scharff¹⁰ have described an immunoglobulin-producing cell that is relevant to both these processes. They showed that the cell line MPC-11 produced two light chains, one apparently normal and the other, a lower molecular-weight fragment, at first thought to contain only constant region sequences. Further studies showed that a derivative of this cell line had lost its ability to produce normal light chains, but that it continued to produce the fragment. In addition, the fragment was encoded in a shorter-than-normal mRNA species¹¹. These studies led to the conclusion that there must be two independently active κ genes in MPC-11 cells and that (at least in this instance) the process of allelic exclusion that affects most antibody-producing cells had been circumvented¹⁰. Subsequent amino acid sequence determinations^{12,13} on an *in vitro* synthesized fragment precursor showed that it consisted of a hydrophobic leader sequence joined directly to the κ constant region and that the sequence corresponding to the variable region (the joined V and J segments) was completely lacking (Fig. 1). Moreover, the hydrophobic leader sequence of the fragment was different from that of the intact light chain produced by MPC-11. Instead, it corresponded to a sequence associated with the VK21 subgroup of κ light chains^{12,14}, an observation that strengthened the notion that each MPC-11 light chain corresponded to a separately expressed allele.

While the simplest explanation for these observations might have been that both chromosomal genes underwent V/J recombination in MPC-11 and that one, the fragment gene, suffered a subsequent deletion of the separately encoded V/J segment, this is not the case. While it might also seem possible that the rearranged gene arose from the normally rearranged gene by a subsequent deletion, this is ruled out by our finding that the normally rearranged chromosome has lost (by deletion associated with V/J recombination) the V-region genes corresponding to those used to form the fragment gene (J.S. and P.L., unpublished data). Since we also found that the fragment gene lacks all J regions, neither gene could serve as precursor for the other and both, therefore, appear to have been derived from independent, allelic germ-line chromosomes. Using DNA cloning, direct sequence analysis and *in situ* hybridization, we show that both alleles have indeed undergone rearrangement in the MPC-11 cell line, but that the fragment gene has undergone a

recombination event that has joined its V region to a site close to the κ constant region, but not associated with any J segment. This joining has apparently been guided by the same palindromic sequence (CACAGTG) that occurs adjacent to all V and J segments thus far sequenced, for the recombination has occurred at the only instance of this palindrome between the J regions and the κ constant-region gene (E. Max, unpublished). Since the J region normally provides the RNA splice site at the 5' end of the intervening sequence between the variable and constant regions, its absence in the fragment gene must prevent the normal RNA splice from occurring. Thus the fragment mRNA must be formed by an aberrant RNA splice involving the hydrophobic leader segment skipping its normal splicing partner, the V segment, and being joined directly to the C region. The nucleotide sequence suggests a mechanism for the aberrant DNA splicing event that involves homologous rather than normal V/J recombination.

In an accompanying study, Choi, Kuehl and Wall¹⁵ show how this mutant gene generates mRNA precursors that follow the required aberrant RNA splicing pathway and that their structure can be reconciled with the structure of the mutant gene described here.

Restriction fragments containing the constant region gene in MPC-11 DNA

Confirmation of the prediction¹⁰ that MPC-11 cells encode two separate and active κ constant-region genes was obtained from *in situ* hybridization experiments in which all the restriction fragments that encode κ constant regions could be identified (Fig. 2). If two constant region genes are present in MPC-11 DNA and have undergone somatic rearrangement, one would expect to find two new restriction fragments each containing a constant region, neither of which should be identical to the fragments observed in embryonic DNA. Cleavage of MPC-11 DNA with the restriction enzyme *Bam*HI produces two such fragments, one 7.7 and the other 3.0 kilobase pairs in length (Fig. 2). The presence of two non-germ-line genes in MPC-11 DNA does not, however, prove that both of these genes are involved in the synthesis of the two light chains expressed in MPC-11 cells. Nor does it allow us to identify the fragment responsible for mutant light-chain production. Other plasmacytoma cells have been shown to contain several non-germ-line genes other than the one which is expressed as a light-chain polypeptide (ref. 7 and E. Max, S. P. Kwan, J. G. S. and P. L., unpublished results). The assignment of the two *Bam*HI restriction fragments to the normal and fragment light-chain genes can

be made by taking advantage of several cell lines derived by Scharff and co-workers from MPC-11 that continue to produce the fragment peptide but no longer produce the normal MPC-11 light chain.

One of these cell lines, NP2, yields DNA containing only the 3-kilobase pair *Bam*HI constant-region containing fragment (Fig. 2). Presumably the 7.7-kilobase pair fragment missing in NP2 but present in the parental line encodes the normal MPC-11 light chain, while the 3.0-kilobase pair *Bam*HI fragment must encode the fragment polypeptide. Apparently the NP2 cells have suffered a deletion of the normal MPC-11 light-chain gene and this deletion accounts for their inability to synthesize this light chain. Interestingly, another cell line, B47L3NP, has also lost the ability to produce the normal MPC-11 light chain, but has retained the 7.7-kilobase pair *Bam*HI fragment (Fig. 2). Obviously the mechanism by which this gene is inactivated is not accounted for by a major deletion, making it a fruitful subject for further investigation.

The MPC-11 fragment gene consists of a variable region and constant region without a J segment

To determine the genetic mechanism that led to the formation of the fragment peptide, we isolated the MPC-11 *Eco*RI DNA fragment that encodes the rearranged fragment gene using preparative agarose gel electrophoresis¹⁷. The proper fragment,

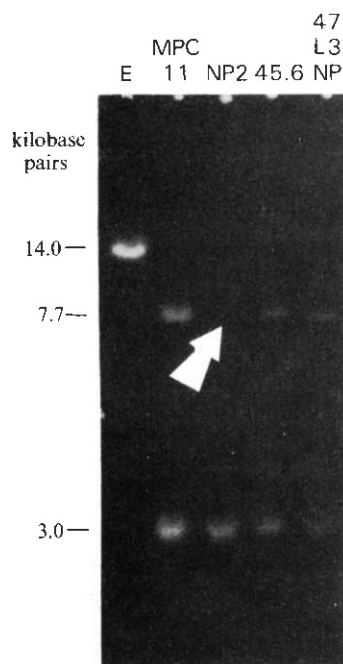


Fig. 2 The constant region containing fragments in DNA derived from embryonic light chain producing and non-producing cells. The DNA from 14-day-old BALB/c embryos (E), MPC-11 and NP2 myeloma tumours, and the 45.6 and B47L3NP (47L3NP) cell lines was restricted with *Bam*HI nuclease, and fractionated on an agarose gel, blotted onto nitrocellulose filters and hybridized to ³²P-labelled constant region probes as described previously^{6,16}. The B47L3NP cell line was derived from 45.6 and no longer secretes a normal κ light chain. NP2 was originally derived from MPC-11 cells as a cell that does not secrete normal MPC-11 light chain, and has since been carried as a myeloma. The constant region probe hybridizes to one fragment of embryonic DNA (14.0 kilobase pairs) but two fragments of MPC-11 DNA (7.7 and 3.0 kilobase pairs). The arrow indicates the absence of the 7.7-kilobase pair fragment in NP2 DNA; we presume that this fragment encodes the MPC-11 normal κ light chain. The 3.0-kilobase pair *Bam*HI fragment derives from the MPC-11 fragment gene.

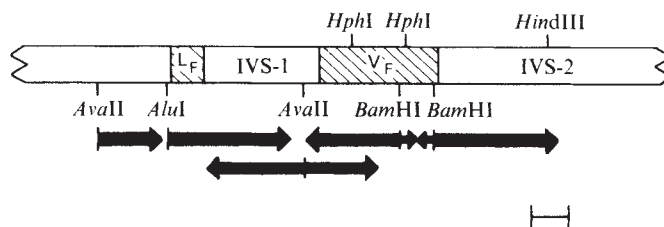


Fig. 3 The restriction fragments used to determine the nucleotide sequence of the MPC-11 fragment gene. A 20-kilobase pair constant region containing *Eco*RI fragment was cloned from MPC-11 DNA into Charon 4A (CH4A MPC-11 fragments) using the procedures described previously⁶. This phage contained a 5-kilobase pair *Hind*III fragment that hybridized to the 2.8-kilobase pair *Hind*III fragment derived from embryonic DNA that contained all of the κ J-region segments and 1.0 kilobase pairs of intervening sequence⁸. The 5-kilobase pair *Hind*III fragment was subcloned into pBR322 (pBR322-MPC-11 fragment) as previously described⁸. Three *Bam*HI fragments (1.8, 0.6 and 0.1 kilobase pairs) were excised from this plasmid and subjected to the nucleotide sequencing procedures described by Maxam and Gilbert¹⁹. The arrows indicate the orientation of the sequences determined and the approximate length of the sequence obtained from each restriction fragment. Scale bar: 0.1 kilobase pairs.

about 20 kilobase pairs long, could be identified because it yielded the expected 3.0-kilobase pair *Bam*HI fragment (Fig. 2). The fragment was cloned in λ vector CH4A, identified and further purified using standard recombinant DNA techniques¹⁸.

Preliminary restriction mapping and hybridization experiments indicated that the fragment gene did not contain any of the κ J-region segments (data not shown). That is, the cloned MPC-11 fragment gene did not hybridize to restriction fragments containing several κ J regions, and it did not contain the *Xba*I, *Sac*I or *Bgl*II restriction sites that are about 1,000 base pairs to the 3' side of the κ J regions (refs 8, 9, and our unpublished results). Thus, it seemed that the MPC-11 fragment gene was formed by a recombination event that occurred about 1,400 base pairs from the κ constant region at a position that contained no J-region genes.

These conclusions could be confirmed and further details revealed by direct sequence determination. Accordingly, an appropriate *Hind*III fragment was sub-cloned from the hybrid phage into pBR322 and the sequence of the relevant region determined using the chemical degradation procedure of Maxam and Gilbert¹⁹ according to the strategy outlined in Fig. 3. The complete sequence of the segment, translated in accordance with the genetic code is shown in Fig. 4.

The fact that the hydrophobic leader segment identified by our sequence was in agreement with the amino acid sequence determined on the fragment peptide^{12,13}, confirmed that we had cloned the correct gene. But the sequence also revealed that the aberrant recombination event that formed the fragment gene had brought in an unexpected, cryptic V-region coding sequence that was not expressed in the fragment polypeptide. In contrast to the situation in normal light-chain genes, the V-region sequence was not adjacent to a J segment.

The MPC-11 fragment gene was formed by an unusual recombination event

Normally a κ light-chain gene is formed by a recombination event that joins one of many κ variable-region genes to one of four J regions to create a gene with the structure indicated in Fig. 1 (normal κ gene). The structure of the MPC-11 fragment gene (Figs 1, 4) is unusual. As noted above, this gene has no J region and lacks part of its intervening sequence. It is, therefore, unlikely that this gene was formed by the site-specific recombination events that normally join V and J segments. Clearly

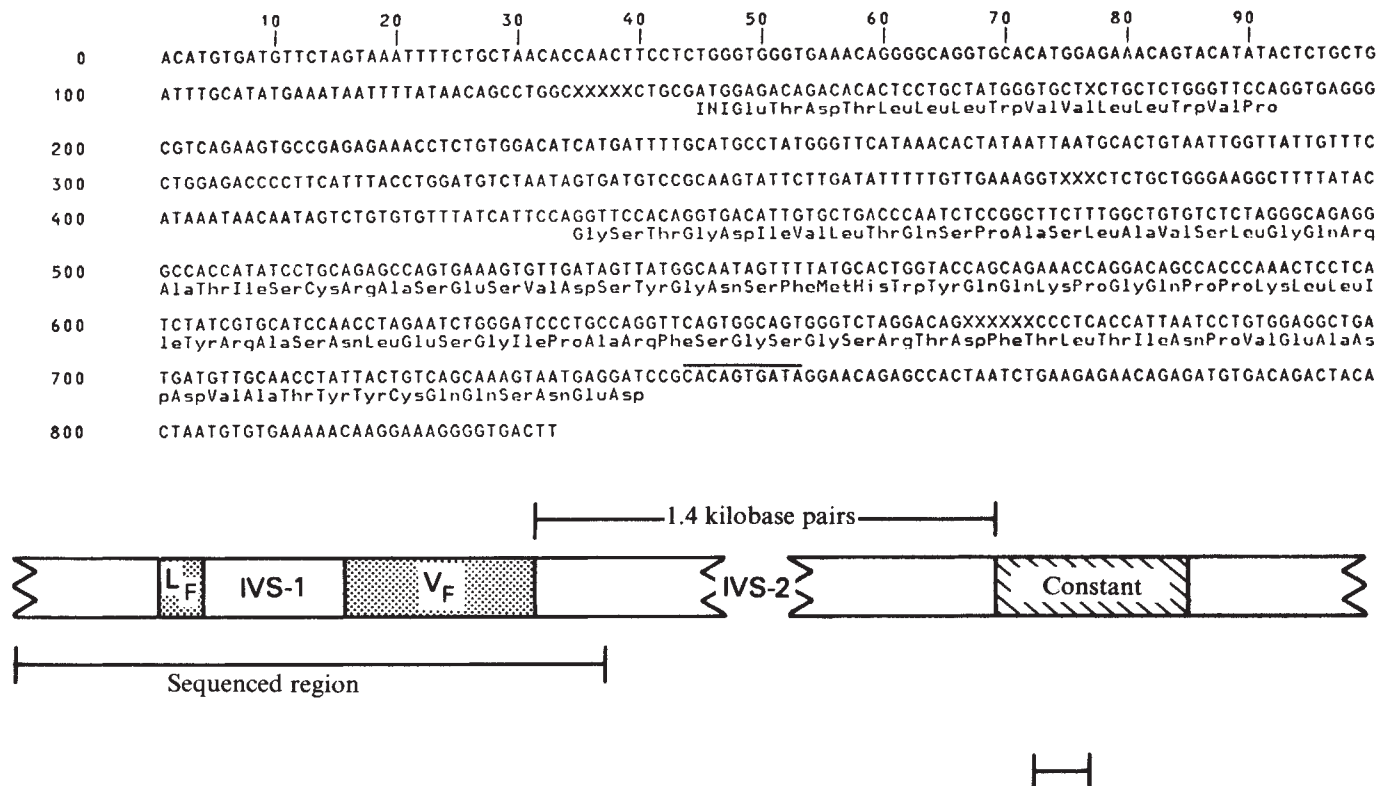


Fig. 4 The nucleotide sequence of the MPC-11 fragment gene. The nucleotide sequence was determined as described in Fig. 3. The portion of the gene that has been sequenced is indicated in the diagram. The leader corresponds to the first amino acid coding block and the variable region to the second coding block. The line over the sequence at the 3' end of the variable region corresponds to the homologous segment found in several κ variable regions and the κ intervening sequence. Scale bar, 0.1 kilobase pairs.

one of the germ-line precursors of this gene must have been the germ-line κ constant-region gene. The fragment containing this gene has been cloned and much of its nucleotide sequence has been determined (refs 8, 9 and E. Max, unpublished data). The 3' end of the MPC-11 fragment gene is identical in sequence to a region about 1.4 kilobase pairs 5' from the K constant-region coding sequence. The other germ-line precursor of MPC-11 fragment gene must be the κ variable-region gene we see here in rearranged form (Fig. 4). From the amino acid translation of its nucleotide sequence, it is clear that this variable region gene is a member of the BALB/c $V_{\kappa}21$ subgroup²⁰. In fact, one of the proteins that falls into this subgroup but is derived from NZB mice ($V_{\kappa}21C$) has an amino acid sequence identical to the variable-region gene found in the MPC-11 fragment gene²¹. It seems that the MPC-11 fragment gene derived its variable region from one member of the BALB/c $V_{\kappa}21$ subgroup.

How was this $V_{\kappa}21$ variable region gene brought to within 1.4 kilobase pairs of the κ constant-region gene into a region that lacks J segments? We and others^{8,9} have suggested that normal V/J gene recombination involves the recognition of a palindromic sequence (CAC^A/TGTG) that is found immediately to the 5' side of all J regions and to the 3' side of all V regions thus far sequenced^{7-9,22,23}. This V/J joining event does not take place by homologous recombination because these palindromic sequences are removed from both the variable and the J region during recombination. Homologous recombination would have preserved one copy of the homologous sequence. The sequence CACAGTG occurs only once between the J- and C-region genes and this is exactly at the site of the recombination event that formed the fragment gene (see below); the palindrome has been preserved in the MPC-11 fragment gene (Fig. 5).

Homology between the germ-line V- and C-region segments, in fact, extends beyond the V/J palindrome. As noted above, the 2.5-kilobase pair intervening sequence that separates the germ-line κ constant-region gene from the J-region segments contains

the palindromic CACAGTG only once at a position about 1.4 kilobase pairs from the C-region gene. By aligning the recombined fragment gene sequence and the germline J/C segment, the point of divergence between the two sequences identifies the point of recombination. When the two sequences were aligned we realized that the recombination event must have taken place within 15 base pairs of the 3' end of the $V_{\kappa}21C$ variable region gene at a point 1.4 kilobase pairs from the C-region gene. The sequence 5'-CACAGTGATA-3' (or a sequence that differs by only a single nucleotide) has been found at the 3' side of the three κ variable-region genes that have thus far been characterized^{7,22}. This sequence is also found at the 3' side of the variable-region gene found in the MPC-11 fragment gene and has also been located at the recombination site in the germ-line J/C fragment. Thus it seems that this fragment gene was formed by homologous recombination taking advantage of

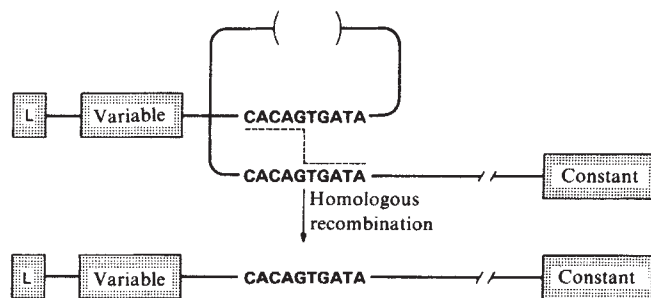
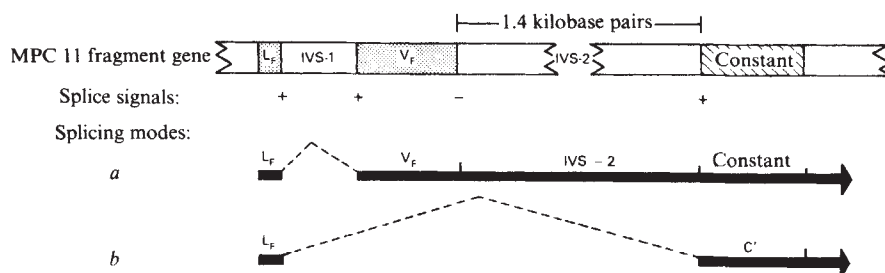


Fig. 5 A model for the homologous recombination event that formed the MPC-11 fragment gene. The data that suggest that this recombination is intrachromosomal and occurs on the non-rearranged chromosome will be presented elsewhere²⁵.

Fig. 6 The two possible splicing modes for RNA transcripts of the MPC-11 fragment gene. By analogy to a normal light-chain gene, the MPC-11 fragment gene must have three of the four normal splice signals (+); the missing splice signal (-) is due to the lack of a J-region gene. These three splice signals could be used in two different modes, to generate two different mRNAs.



sequences that are found on both the germ-line variable-region and κ constant-region/J-region intervening sequences. The event is shown diagrammatically in Fig. 5. This finding further suggests that this position might behave as a hot spot for aberrant light-chain gene formation. Indeed, similar genes may account for short mRNA species seen in a number of plasmacytoma cell lines²⁴.

Forbidden splicing of the precursor RNA is required to form the fragment peptide

A comparison of the structures of the fragment peptide and its corresponding gene presents something of a paradox in that the V-region sequence is present in the gene, but absent in the fragment polypeptide (Fig. 1). Normally, the V-region coding sequence is joined to a J segment that must contribute, in addition to a 13-codon long segment of DNA, a signal for RNA splicing so that the V/J segment can be joined to the constant-region sequence in the mature mRNA. In the case of the fragment gene, there is no J segment and therefore apparently no signal for joining V and C regions into a continuous mRNA. The theoretical splicing alternatives for the fragment gene transcript are shown diagrammatically in Fig. 6. A normal splice pair exists between the leader and V-region segments (Fig. 6a), but this splice would lead to the formation of leader-variable region peptides coupled to the translation product derived from the intervening sequence. No such peptides have been sought or found. The only way the mRNA corresponding to the fragment peptide could be formed is by joining the leader sequence directly to the constant-region sequence (Fig. 6b), skipping the normal leader/V splice. Formation of the fragment peptide must, therefore, involve not only an abnormal DNA recombination event but also an abnormal RNA splice.

Ordinarily an RNA splice that produces a mutant protein would be expected to be quite rare. Errors in splicing occurring in even a small proportion of splice segments would seem disadvantageous in a large gene containing many splice junctions since they would seriously reduce the yield of normal mRNA. Evidently some feature of the transcript must specify the splicing order, assuring that proper donors and acceptors are matched and joined. The fragment light chain gene evidently lacks this control. The simple notion that processing can only occur by splicing next to adjacent donors and acceptors is ruled out by this mutant gene, because the existence of the MPC-11 fragment proves that the constant region competes successfully

with the variable region for joining to the leader sequence (fragment mRNAs constitute roughly 20% of normal mRNA levels¹¹). On the other hand, it is possible to imagine that forbidden splices occur only when there is an imbalance between donors and acceptors. That is, perhaps a preference for adjacent donor/acceptor joining normally reduces the availability of unmatched partners and provides sufficient specificity to guide normal splicing. The difficulty with this notion, while accounting quite well for the events that occur in this mutant light-chain gene, is that it does not satisfactorily account for the variation in splicing observed in several viral mRNA precursors in which alternative splice pathways are frequently used²⁶⁻²⁸.

Violation of allelic exclusion: activation of light-chain promoters

Kuehl and Scharff and their co-workers¹⁰ concluded that, in contrast to the usual situation in antibody-producing cells, two genes must be active in the MPC-11 cell. The basis by which allelic exclusion is avoided in MPC-11 can be discerned by taking into account the structure of these genes and the recent observation by Perry *et al.*²⁴ that both constant-region genes, germ line and rearranged, are actively transcribed in many plasmacytomas. Most antibody-producing cells contain only one gene that has undergone V/J recombination so as to produce a normal light chain. Notwithstanding the fact that both genes are transcribed, only the properly rearranged one will produce a light-chain product. In the case of MPC-11, one normal light-chain gene has presumably been formed (Fig. 2), but the fragment gene, though abnormally formed, is actively transcribed and has sufficient elements of structure so as to produce a mutant light chain. This observation, taken together with those of Perry *et al.*²⁴, suggest to us that a higher order of control activates regions of a chromosome and that promoters that happen to fall within this region will be actively transcribed. For example, all V-region sequences might carry their own promoters which remain inactive until moved into an open region of a B-cell chromosome. Such a region obviously surrounds the C-region locus. The availability of many cloned V-region segments suggests that the existence of multiple, independent V-region promoters can be tested experimentally.

We thank Drs Michael Kuehl and Matthew Scharff for their gifts of cells used in these studies and for their helpful advice, Barbara Norman and Marion Nau for their participation in these studies and Terri Broderick for her assistance in the preparation of the manuscript.

Received 4 April; accepted 5 May 1980.

- Dreyer, W. J. & Bennett, J. C. *Proc. natn. Acad. Sci. U.S.A.* **54**, 861-868 (1965).
- Hozumi, N. & Tonegawa, S. *Proc. natn. Acad. Sci. U.S.A.* **73**, 3628-3632 (1976).
- Rabbitts, T. H. & Forster, A. *Cell* **13**, 319-327 (1978).
- Seidman, J. G. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **75**, 3881-3885 (1978).
- Brack, G., Hiram, M., Lenhard-Schuller, R. & Tonegawa, S. *Cell* **15**, 1-14 (1978).
- Seidman, J. G. & Leder, P. *Nature* **276**, 790-795 (1978).
- Seidman, J. G., Max, E. E. & Leder, P. *Nature* **280**, 370-375 (1979).
- Max, E. E., Seidman, J. G. & Leder, P. *Proc. natn. Acad. Sci. U.S.A.* **76**, 3450-3454 (1979).
- Sakano, H., Hüppi, K., Heinrich, G. & Tonegawa, S. *Nature* **280**, 288-294 (1979).
- Kuehl, W. M. & Scharff, M. D. *J. molec. Biol.* **89**, 409-421 (1974).
- Kuehl, W. M. *et al.* *Cell* **5**, 139-147 (1975).
- Rose, S. M., Kuehl, W. M. & Smith, G. P. *Cell* **12**, 453-462 (1977).
- Burstein, Y. & Schechter, I. *Proc. natn. Acad. Sci. U.S.A.* **74**, 716-720 (1977).

- Burstein, Y. & Schechter, I. *Biochem. J.* **157**, 145-151 (1976).
- Choi, E., Kuehl, M. & Wall, R. *Nature* **286**, 776-779 (1980).
- Southern, E. M. *J. molec. Biol.* **98**, 503-517 (1975).
- Polisky, F., Edgell, M., Seidman, J. & Leder, P. *Analyt. Biochem.* **87**, 397-410 (1978).
- Tiemeier, D., Tilghman, S. & Leder, P. *Gene* **2**, 173-191 (1977).
- Maxam, A. & Gilbert, W. *Proc. natn. Acad. Sci. U.S.A.* **74**, 560-564 (1977).
- McKean, D., Potter, M. & Hood, L. *Biochemistry* **12**, 794-759 (1973).
- Weigert, M., Gatmaitan, L., Loh, E., Schilling, J. & Hood, J. *Nature* **276**, 785-790 (1978).
- Seidman, J. G., Leder, A., Nau, M., Norman, B. & Leder, P. *Science* **202**, 11-17 (1978).
- Bernard, O., Hozumi, N. & Tonegawa, S. *Cell* **15**, 1133-1144 (1978).
- Perry, R. P. *et al.* *Proc. natn. Acad. Sci. U.S.A.* (in the press).
- Seidman, J. *et al.* *Proc. natn. Acad. Sci. U.S.A.* (in the press).
- Berk, A. & Sharp, P. A. *Cell* **14**, 695-711 (1978).
- Kitchingman, G. R., Lai, S. P. & Westphal, H. *Proc. natn. Acad. Sci. U.S.A.* **74**, 4392-4395 (1977).
- Gelinas, R. E. & Roberts, R. J. *Cell* **11**, 533-544 (1977).

LETTERS

 γ -Ray burst observations from the UCB/LASL experiment on ISEE-3

W. D. Evans, R. W. Klebesadel, J. G. Laros & J. Terrell

University of California, Los Alamos Scientific Laboratory,
Los Alamos, New Mexico 87545

S. R. Kane

Space Sciences Laboratory, University of California,
Berkeley, California

Since the discovery of the intense bursts of γ rays by Klebesadel *et al.*¹ in 1973, little progress has been made in the identification of the sources of the bursts. The initial observations provided directions to several sources with sufficient accuracy to establish the approximate isotropic distribution of the sources², but not to provide a basis for association with any known objects. γ -ray burst astronomy entered a new phase in 1976 with the launch of Helios 2 and Solrad 11 containing the first experiments designed expressly for that purpose^{3,4}. It then became possible to derive several error boxes that were very narrow in one dimension^{3,5}. Using these data it was determined that the bursts are not associated with any type of object that has previously been recognized as a high energy astrophysical source—for example, X-ray sources. In 1978 a true interplanetary network of burst detectors was established with the launch of instruments on Pioneer Venus Orbiter⁶, ISEE 3⁷⁻⁹ and Venera 11 and 12^{10,11}. With baseline separations that can exceed 1 AU, this network provides a basis for the location of the γ burst sources with an accuracy sufficient for optical identifications. We briefly describe here one of these instruments, a joint Berkeley/Los Alamos experiment on the ISEE 3 spacecraft, and present some initial results from that experiment. These results are currently being combined with those from the other experiments to obtain accurate burst locations to be reported elsewhere.

The University of California X-Ray Spectrometer⁷ was originally designed to monitor solar emission continuously over the energy range 4.8–14 keV (1.2 cm² Xe/CO₂ proportional counter) and 12–1,264 keV (22 cm² NaI(Tl) scintillation detector). Additional electronic instrumentation, provided by the Los Alamos Scientific Laboratory (LASL), was later incorporated into the experiment to record the data from the intense, short lived γ -bursts at a very high rate and with good timing accuracy. This addition consisted of a data handling system that accepts photon counts from the higher energy PHA channels of the scintillation detector and monitors the counting rate for rapid increases that indicate either a γ burst or a solar flare. When a transient event of either origin occurs, high time resolution data are accumulated in a 36 kbit solid state memory for later transmission at a much reduced rate. During a γ burst the summed counting rate from six PHA channels covering the energy range 136–1,264 keV is recorded at the highest time resolution. For counting rates below 1,365 s⁻¹ the accumulations are made over fixed intervals of 11.72 ms. At higher rates the operation automatically changes to a 'time-to-spill' mode wherein the resolution improves linearly with counting rate to an ultimate 0.24 ms, while the statistical significance of each accumulation remains constant. Six-channel energy spectra covering the same 136–1,264 keV range are accumulated simultaneously at a factor of 12 coarser time resolution. None of the bursts observed to date have challenged the counting rate capacity or ultimate time resolution of the instrument. (The extremely intense 5 March 1979 event was largely blocked by the spacecraft.)

The ISEE 3 spacecraft was launched in August 1978 and was placed into a heliocentric orbit so that it remains approximately on the Earth-Sun line (at the inner Lagrangian point) at a geocentric distance of 0.01 AU (5 light s). Between launch and 19 April 1979, 12 γ bursts, each confirmed by other instruments, have been detected and recorded by the UCB/LASL experiment. Bursts known to have occurred on 24 November 1978 and 13 January 1979 were missed due to incomplete (~90%) satellite tracking coverage. Virtually all spurious events that we have seen here were caused by statistical fluctuations in the background counting rate, as opposed to cosmic ray-induced noise or other non-statistical sources. These spurious events can occur as frequently as ~1 per day, depending on the trigger criteria in effect. As the criteria are selectable by command, we have sacrificed some sensitivity to eliminate the spurious events during most of the mission. Table 1 lists the dates and times of the confirmed γ burst events and their approximate peak intensities (defined as the highest 0.14 s wide time-history bin) and total integrated fluxes. The latter two quantities have not been corrected for detector efficiency or spacecraft shielding and scattering. Thus, the intensity entries in Table 1 are useful only as a very rough guide and as an indication of the sensitivity and dynamic range of the instrument. For example, it can be inferred that the sensitivity limit of the instrument is ~2 photons cm⁻² ($\leq 10^{-6}$ erg cm⁻²). Of course, more intense bursts might be missed because of their spectra, directions, or time histories.

Figure 1 shows samples of the temporal and spectral data obtained so far. We do not discuss the singular event of 5 March 1979 as it has been treated extensively elsewhere (see, for example, refs 12–14). The 4 November 1978 event (Fig. 1), though unusually intense, exhibits most of the classical γ -burst characteristics such as fast rise and fall times, multiple peaked structure, and near background rates between peaks. As can be seen qualitatively in Fig. 1A*b*, a χ^2 analysis of the spectral information showed no evidence for spectral variability, to within the statistical uncertainties. In view of the complex time history, this spectral constancy above 130 keV seems surprising, and must bear critically on the emission mechanism for this burst. The Apollo 16 γ burst^{15–17} and others⁴ also featured complex structures combined with no detectable spectral variability. Finally, an intriguing pattern can be seen in the time history of this event. Namely, the bulk of the emission is concentrated in three double peaks, with ~9 s between doublets and with the components of the doublets separated in time by amounts proportional to their intensities. While we make no claims regarding the significance of this possible pattern, we point it out as the type of information that can be extracted from continuous, high time resolution γ -burst data records. Clearly, the short-lived nature of the bursts means that a collection of

Table 1 γ bursts observed by the UCB/LASL experiment on ISEE 3

Date	Time (UT)	Intensity (photons cm ⁻² s ⁻¹)	Flux (photons cm ⁻²)
14 September 1978	16.42.08	*	*
18 September 1978	19.49.33	7	7.5
21 September 1978	03.55.56	9.2	7.5
19 October 1978	18.26.46	2.8	7.5
4 November 1978	16.17.49	29	80
15 November 1978	21.07.25	2.5	3
19 November 1978	09.26.58	28	65
5 March 1979	15.52.05	74	16
7 March 1979	22.18.51	6.5	27
31 March 1979	21.09.36	3	2.8
18 April 1979	07.41.02	8.8	9.3
19 April 1979	16.13.43	3.8	7.8

* Data were lost during the intense part of the burst.

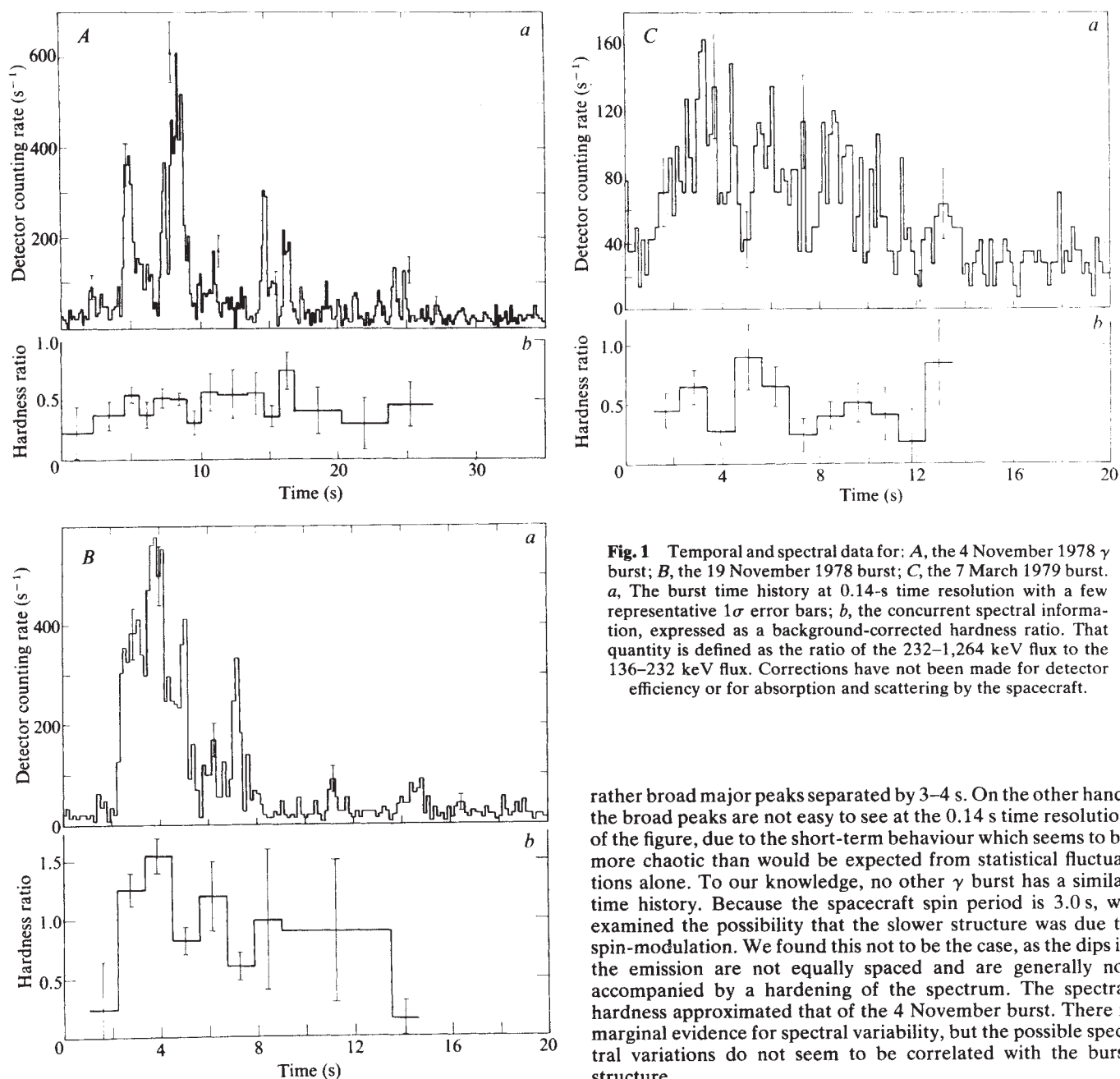


Fig. 1 Temporal and spectral data for: *A*, the 4 November 1978 γ burst; *B*, the 19 November 1978 burst; *C*, the 7 March 1979 burst. *a*, The burst time history at 0.14-s time resolution with a few representative 1σ error bars; *b*, the concurrent spectral information, expressed as a background-corrected hardness ratio. That quantity is defined as the ratio of the 232–1,264 keV flux to the 136–232 keV flux. Corrections have not been made for detector efficiency or for absorption and scattering by the spacecraft.

such information is needed before the significance of this or other possible patterns can be assessed.

The 19 November 1978 burst (Fig. 1*B*) was also one of the most intense ever recorded. As pointed out by Evans *et al.*⁶, its time history is atypical in that after the initial sharp rise the counting rate remained at a very high level for an unusually long time. Alternatively, the initial 3 s of the burst might appear to consist of a blend of several sharp peaks that are atypically close together. Also, the average spectrum of the burst is unusually hard, as can be seen by comparing Fig. 1*B* with Fig. 1*A* and *C*. Finally, there is clear evidence for spectral variability correlated with the burst structure. Although the counting statistics do not allow a precise correlation, there is no doubt that the sharp spikes seen in Fig. 1*Aa* at ~ 5 and at ~ 7 s have softer spectra than the early part of the burst, and this is probably true also for the spikes at 11 and 14 s. Of course, the spectral hardness may simply be decreasing monotonically.

The 7 March 1979 burst (Fig. 1*C*), though not particularly intense at any given instant, had a fairly large total flux and an unusual time history. It is possible to discern in Fig. 1*Ca* several

rather broad major peaks separated by 3–4 s. On the other hand, the broad peaks are not easy to see at the 0.14 s time resolution of the figure, due to the short-term behaviour which seems to be more chaotic than would be expected from statistical fluctuations alone. To our knowledge, no other γ burst has a similar time history. Because the spacecraft spin period is 3.0 s, we examined the possibility that the slower structure was due to spin-modulation. We found this not to be the case, as the dips in the emission are not equally spaced and are generally not accompanied by a hardening of the spectrum. The spectral hardness approximated that of the 4 November burst. There is marginal evidence for spectral variability, but the possible spectral variations do not seem to be correlated with the burst structure.

To obtain a more quantitative comparison of the burst time histories, a Fourier analysis program was carried out. The methods used were not entirely the standard techniques, but have been used or discussed previously^{14,18,19}. Each time series consisted of 35.86 s of data in 3,060 time bins each 11.72 ms wide. The data were subjected only to linear detrending; no grouping or smoothing was used. The spectra were examined up to the Nyquist frequency of 42.67 Hz, but only to region below 8 Hz is shown in Fig. 2, since there is little significant higher frequency information. The results generally confirm our previous qualitative statements on the burst time histories, and also point out some less obvious features. The power spectra of the two November 1978 events both show significant peaks in the 0.1–1 Hz region, indicating a degree of regularity in their time histories. One curious pattern that was pointed out above in the 4 November 1978 burst, probably produced the spectral peak at 0.1 Hz. The 0.3 Hz spectral peak is most likely due to the ~ 3 s separation of the spikes seen in Fig. 1*A* at 2, 5, 8, 11, and 14 s. Also, further examination of the 19 November time history shows that the several count rate peaks are separated by ~ 1.1 s or a multiple thereof, giving rise to the 0.9 Hz spectral peak. The very broad peak below 0.3 Hz reflects the duration of the main

part of the γ burst, but may also be due in part to the widely separated counting spikes near the end of the burst. On the other hand, the power spectrum of the 7 March 1979 event approximates that of white noise due to counting fluctuations, except for the low-frequency power below ~ 0.5 Hz. Thus, the chaotic appearance of the 7 March time history (Fig. 1Ca) is either not statistically significant, or contributes only a small amount of power spread approximately uniformly over a broad frequency band. The spiky appearance of this spectrum is due mainly to the greatly magnified vertical scale relative to the other power spectra shown. Of the three bursts, only the 19 November event

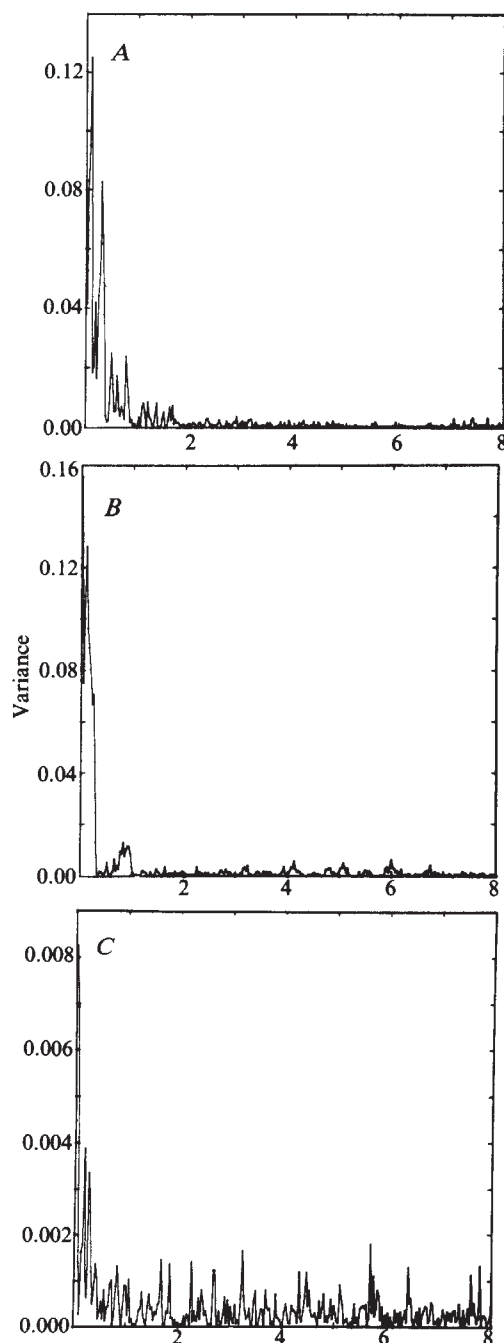


Fig. 2 Power spectra for the three γ -ray bursts shown in Fig. 1. A, 4 November 1978; B, 19 November 1978; C, 7 March 1979. The ordinate is variance per Fourier harmonic. Each time series consisted of 3,060 accumulations at time intervals of 11.72 ms for a total time of 35.86 s. The data have been detrended before analysis in each case. No grouping or smoothing has been used. Only the lower portion of the frequency spectrum, up to 8 Hz, is shown (the Nyquist frequency is 42.67 Hz).

clearly shows excess power at higher frequencies, even beyond 8 Hz. This high frequency power undoubtedly results from the sharp counting rate transitions evident, though not optimally displayed, in Fig. 1B.

This temporal and spectral information from the UCB/LASL γ burst experiment on ISEE 3 indicates that the three bursts studied had comparable integrated fluxes, but showed significant differences in their power spectra as well as in the qualitative appearances of their time histories. There are also important differences in their average energy spectra, and in their degree of spectral variability. Thus, the present examples show that the γ -burst phenomenon is diverse, and that the sources of the bursts constitute either a very loose class or possibly more than one class of objects. This fact must be taken into account in analyses, for example, the interpretation of $\log N/\log S$ curves, which assume a 'standard' γ burst object.

We thank Earl Tech and other members of the data analysis section of the LASL high altitude physics group for assistance, and Bill Aiello and Henry Primbsch for their work on the experiment hardware. This research was supported by NASA under contract NAS5-22307 and by the US Department of Energy.

Received 15 January; accepted 16 June 1980.

1. Klebesadel, R. W., Strong, I. B. & Olson, R. A. *Astrophys. J. Lett.* **182**, L85 (1973).
2. Strong, I. B., Klebesadel, R. W. & Olson, R. A. *Astrophys. J. Lett.* **188**, L1 (1974).
3. Cline, T. L. *et al. Astrophys. J. Lett.* **229**, L47 (1979).
4. Laros, J. G., Evans, W. D., Klebesadel, R. W., Olson, R. A. & Spalding, R. E. *Nature* **267**, 131 (1977).
5. Cline, T. L. *et al. Astrophys. J. Lett.* **232**, L1 (1979).
6. Evans, W. D. *et al. Science* **205**, 119 (1979).
7. Anderson, K. A. *et al. IEEE Trans. Geosci. Elect.* **GE-16**, 157 (1978).
8. Hovestadt, D. *et al. IEEE Trans. Geosci. Elect.* **GE-16**, 166 (1978).
9. Cline, T. L., Gloeckler, G., Hovestadt, D. & Teegarden, B. J. *IEEE Trans. Geosci. Elect.* **GE-16**, 173 (1978).
10. Chambon, G. *et al. Proc. 1979 COSPAR Conf.* (in the press).
11. Mazets, E. P. *et al. Astr. Zhurnal-Pis'ma* **5**, 163 (1979).
12. Evans, W. D. *et al. Astrophys. J. Lett.* **237**, L7 (1980).
13. Cline, T. L. *et al. Astrophys. J. Lett.* **237**, L1 (1980).
14. Terrell, J., Evans, W. D., Klebesadel, R. W. & Laros, J. G. *Nature* **285**, 383 (1980).
15. Metzger, A. E., Parker, R. H., Gilman, D., Peterson, L. E. & Trombka, J. I. *Astrophys. J. Lett.* **194**, L19 (1974).
16. Trombka, J. I. *et al. Astrophys. J. Lett.* **194**, L27 (1974).
17. Gilman, D., Metzger, A. E., Parker, R. H., Evans, L. G. & Trombka, J. I. *Astrophys. J.* **236**, 951 (1980).
18. Terrell, J. & Olsen, K. H. *Astrophys. J.* **161**, 399-413 (1970).
19. Terrell, N. J. *Astrophys. J. Lett.* **174**, L35-L41 (1972).

A high sensitivity determination of the hard X-ray spectrum of Sco X-1

R. E. Rothschild, D. E. Gruber, F. K. Knight, P. L. Nolan & Y. Soong

Center for Astrophysics and Space Sciences, University of California at San Diego, La Jolla, California 92093

A. M. Levine, F. A. Primini, W. A. Wheaton* & W. H. G. Lewin

Department of Physics and Center for Space Research, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

The results of hard X-ray observations of Sco X-1 by HEAO 1 and OSO 7 are reported here. We have found that the X-ray spectrum between 20 and 70 keV is consistent with thermal bremsstrahlung from a hot plasma¹ at a temperature of 5.15 ± 0.05 keV. Observations with the instruments of HEAO 1 on 6-7 September 1978, and with those of OSO 7 during 1972, have not confirmed reports² that the bremsstrahlung spectrum is accompanied by an extra component at energies greater than 40 keV—indeed, our observations yield a two-standard deviation upper limit to the flux in this region of 6.4×10^{-6} photons $\text{cm}^{-2} \text{s}^{-1} \text{keV}^{-1}$.

* Present address: Jet Propulsion Laboratory, Pasadena, California.

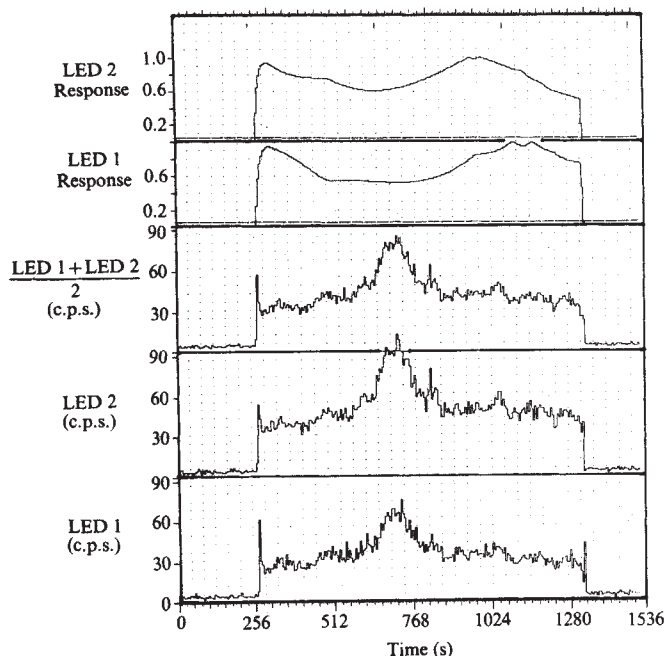


Fig. 1 The aspect-corrected counting rates for the two low energy detectors (LED) aboard HEAO 1 during one 17-min exposure to Sco X-1 on 7 September 1978 covering the energy range 20–220 keV. Also shown is the mean of the two sets of rates and the collimator response to Sco X-1 for each detector. Time 0.000 corresponds to 1 h 8 min 22 s UT. c.p.s., counts per second.

The UCSD/MIT high energy X-ray and low energy γ -ray instrument³ aboard HEAO 1 consisted of seven NaI(Tl)/CsI(Na) phoswich scintillation detectors covering various energy ranges from 10 keV to 10 MeV that were actively shielded by 4–10 cm of CsI. Two of these detectors, of $\approx 100 \text{ cm}^2$ each, sensitive over the range 20–220 keV, were collimated by Sn/Cu slats to $1.5^\circ \times 20^\circ$ FWHM. An observation of Sco X-1 was conducted on 6–7 September 1978 by alternating between source and background regions separated by 2° with 17-min segments for each. This length of time was chosen to sample uniformly, on the average, regions of different particle-induced background due to varying cosmic-ray cutoff throughout the orbit. In this way an important systematic uncertainty in background subtraction could be minimized.

The HEAO 1 observation provided 17 separate exposures to Sco X-1; of these, one displayed pronounced variability, that is an X-ray flare which is shown in Fig. 1. In addition to the 2-min flaring by a factor of ≈ 2 , this exposure differs from the others by having about twice the average intensity off the flare of the others. As spectral shape may vary with intensity¹, the exposure containing the flare was not included in the accumulation of the mean spectrum. It was, instead, studied in detail for spectral changes during this period of both enhanced intensity and activity.

The net counts spectrum was fitted by folding various analytical models through the instrumental response, which included photopeak efficiency, non-linear pulse height to energy conversion, resolution, and K-escape effects, to form a model counts spectrum. This was then compared with the net counts spectrum, and best-fit parameters were determined by minimizing χ^2 . Spectrum dependent efficiencies were calculated and divided into the net counts spectrum to form the inferred mean incident photon spectrum, which is shown in Fig. 2.

The mean spectrum was fit by both a thermal bremsstrahlung model, including a Gaunt factor, and the same model modified to account for Compton interactions in a hot, fully ionized cloud surrounding the X-ray source⁴. The resultant best-fit temperatures were $5.15 \pm 0.05 \text{ keV}$ and $3.24 \pm 0.05 \text{ keV}$, respectively,

and the latter fit was insensitive to optical depth to electron scattering in the range $\tau = 5$ –20. The best-fit temperatures are in agreement with earlier results at lower energies^{1,4}. The inferred mean incident flux and the best-fit ($kT = 5.15 \text{ keV}$) thermal bremsstrahlung model are shown in Fig. 2 together with the inferred flux from the flare alone. An attempt was then made to test for the presence of a high energy excess in the mean spectrum with respect to either thermal model. Assuming a power law form for the excess ($dN/dE \propto E^{-1}$), none was detected at 95% confidence and the 2σ upper limit to any excess is $6.4 \times 10^{-6} \text{ photons cm}^{-2} \text{ s}^{-1} \text{ keV}^{-1}$, or $0.56 \mu\text{Jy}$, in the range 72–222 keV.

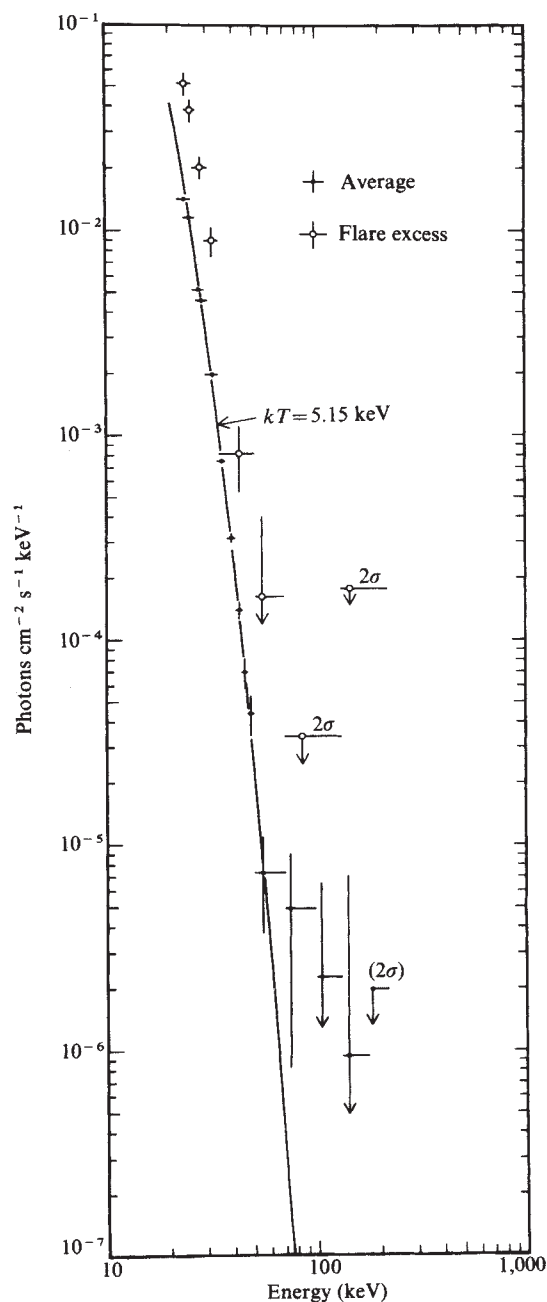


Fig. 2 The inferred incident spectrum of Sco X-1 from 20 to 220 keV from the observations by the HEAO 1 satellite beginning on 6 September 1978. ●, The flux averaged over the entire 12-h observation with the exception of one 17-min interval that contained a flare of 2-min duration (see Fig. 1). ○, The flux from the flare alone with the remainder of the 17-min exposure used as background. All error bars are 1σ unless otherwise noted.

The study of the flare exposure was divided into the flare alone (from 655 to 778 s in Fig. 1) and the remainder of the exposure (from 256 to 655 s and 778 to 1,331 s in Fig. 1). The spectrum of the flare alone, which is also shown in Fig. 2, is the net flux above the remainder of this exposure, and can be compared to the mean spectrum of the other exposures to determine if the same process could have caused both spectra. Acceptable fits to the flare alone are possible with best-fit values for kT of 6.1 ± 0.5 keV and 3.5 ± 0.5 keV for the two thermal models, respectively, or with a power law index of 6.6 ± 0.2 . The narrow energy range over which the flare was observed, however, precludes the selection of one model over another, but the data are consistent with a thermal origin for the flare. No evidence for a high energy excess above the thermal model was seen in the flare, with the 2σ upper limit being 1.4×10^{-4} photons $\text{cm}^{-2} \text{s}^{-1} \text{keV}^{-1}$, or 11 μJy , from 52 to 222 keV.

For the remainder of the flare exposure, the best-fit values for kT of this enhanced activity were 4.9 ± 0.2 and 3.1 ± 0.1 keV, respectively, for the two thermal models mentioned above and the 2σ upper limit to any excess flux was 3.4×10^{-5} photons $\text{cm}^{-2} \text{s}^{-1} \text{keV}^{-1}$, or 3.1 μJy , from 72 to 222 keV.

The UCSD OSO 7 SKYMAP⁵ contains 280 s of live time to Sco X-1 from scanning in mid (May–June) and late (November–December) 1972. These data come from a 64 cm^2 NaI(Tl) detector sensitive from 5 to 500 keV, that was actively shielded by CsI(Na) and had a field of view of 6.5° FWHM. This observation differs from that of HEAO 1 in that data were accumulated over two 18-day periods separated by six months, and thus represents more of a long-term average compared with the brief observation by HEAO 1.

The OSO 7 observation of Sco X-1 was fitted in a similar manner. The fitting process yielded a best-fit temperature of 6.5 ± 0.5 keV for the thermal bremsstrahlung model and 3.2 ± 0.2 keV for the Compton modified thermal model. The ≈ 1.5 keV higher temperature as compared to the HEAO 1 result is consistent with the range of temperatures reported over the past several years for Sco X-1 (ref. 1). As in the HEAO 1 measurements, no excess above the thermal spectrum was observed, and the 2σ upper limit to the flux from 60 to 208 keV was 6.2×10^{-5} photons $\text{cm}^{-2} \text{s}^{-1} \text{keV}^{-1}$, or 5.4 μJy .

Over the past decade and a half, only four observations claim 3σ or greater excesses above a thermal model with fluxes of a few 10^{-4} photons $\text{cm}^{-2} \text{s}^{-1} \text{keV}^{-1}$ in the 50–100 keV range^{6–9}. All four of these have systematic uncertainties, such as possible source confusion, varying background counting rates, and broad instrumental resolution, that reduce the significance of their results. If, indeed, this excess is time variable, however, the magnitude of the change must be more than an order of magnitude at 100 keV.

Models of the Sco X-1 binary system generally include mass flow onto a compact object surrounded by an accretion disk. Both thermal bremsstrahlung¹⁰ and power law¹¹ spectral shapes have been predicted for the X-ray data. The observed temperature for thermal models will not be a true measure of the plasma temperature as various proposed processes might alter the spectral shape. A high-temperature plasma ($kT \approx 75$ keV) could be degraded by Compton interactions to form a spectrum with a lower measured temperature¹². On the other hand, an intense soft X-ray flux ($kT \approx 3$ keV) impinging on the plasma material could undergo Compton scattering to higher energies with either a resultant higher measured temperature⁴ or possibly the appearance of a high energy excess¹⁰ or both. The absence of an observed high energy excess indicates that the mean free path for inverse Compton scattering must be either so low that the scattered flux is below instrument sensitivity, or so high that the reprocessed spectrum resembles an exponential of higher measured temperature. A third possibility is that kT is not high enough to scatter photons to greater than 40 keV.

The nature of the collapsed object powering the X-ray emission in Sco X-1 is uncertain. The minimum mean luminosity of this observation is 1.3×10^{36} erg s^{-1} , using 270 pc as a minimum distance¹³. This places it near the limiting luminosity for spheri-

cal accretion onto a white dwarf¹⁴. Add to this the factor of 5 overall increase in the flare, and it seems quite unlikely that spherical accretion onto a white dwarf could be responsible. On the other hand, accretion onto a neutron star could easily generate such luminosity. Non-spherical accretion onto white dwarfs, however, can also generate luminosities in the 10^{37} erg s^{-1} range that may be observable¹ (N. D. Kylafis, personal communication), and, therefore, allow the white dwarf hypothesis.

We thank R. M. Pelling, J. L. Matteson, W. A. Baity, H. Hudson, B. Dennis, M. Coe and N. Kylafis for their discussions, and R. Pegg and G. Huszar for computer support. This work was supported by NASA contract NAS-8-27974.

Received 10 April; accepted 26 June 1980.

1. Miyamoto, C. & Matsuoka, M. *Space Sci. Rev.* **20**, 687–755 (1977).
2. Agrawal, P. C. *et al. Astrophys. Space Sci.* **10**, 500–507 (1971).
3. Matteson, J. L. *Proc. AIAA 16th Aerospace Sci. Meet.* Huntsville, Alabama, paper 78-35 (1978).
4. Lamb, P. & Sanford, P. W. *Mon. Not. R. astr. Soc.* **188**, 555–563 (1979).
5. Rothschild, R. E., Baity, W. A., Mushotzky, R. F., Wheaton, W. A. & Peterson, L. E. (*COSPAR*) *X-Ray Astronomy* (eds Baity, W. A. & Peterson, L. E.) 369–372 (Pergamon, Oxford, 1979).
6. Haymes, R. C., Harnden, Jr, F. R., Johnson III, W. N. & Prichard, H. M. *Astrophys. J. Lett.* **172**, L47–49 (1972).
7. Riegler, G. R., Boldt, E. & Serlemitsos, P. *Nature* **226**, 1041–1043 (1970).
8. Peterson, L. E. & Jacobson, A. S. *Astrophys. J.* **145**, 962–965 (1966).
9. Greenhill, J. G., Coe, M. J., Bell-Burnell, S. J., Strong, K. T. & Carpenter, G. F. *Mon. Not. R. astr. Soc.* **189**, 563–570 (1979).
10. Shakura, N. I. *Soviet Astr.* **16**, 756–762 (1973).
11. Pringle, J. E. & Rees, M. J. *Astr. Astrophys.* **21**, 1–9 (1972).
12. Kylafis, N. D. & Lamb, D. Q. *Astrophys. J. Lett.* **228**, L105–110 (1979).
13. Bahcall, J. N. *Physics and Astrophysics of Neutron Stars and Black Holes* (eds Giacconi, R. and Ruffini, R.) 63–110 (North-Holland, Amsterdam, 1978).
14. Kylafis, N. D. thesis, Univ. Illinois (1978).

Solar-flare produced ^3He in lunar samples

M. N. Rao & T. R. Venkatesan

Physical Research Laboratory, Ahmedabad 380 009, India

Lunar rocks and soils undergo long-term irradiation by solar and galactic cosmic rays (SCR and GCR) leading to compositionally-distinct spallation spectra for noble gases^{1,2} and leaving behind fossil records characteristic of different irradiation features. Because the SCR effects are restricted to the first few millimetres of rocks and soils in contrast to GCR, various studies based on solar flare tracks^{3,4}, thermoluminescence⁵, and radionuclides^{3,6} over the past decade have aimed to decipher these fossil records bearing the imprints of the low (SCR) and high (GCR) energy irradiation phenomena. These studies have greatly helped our understanding of the energy spectra and fluxes of charged particles streaming out of the Sun. We have already reported quantitative studies, aimed at isolating the SCR and GCR produced Ne and Ar isotopes in several lunar soils and rocks^{1,2} and have calculated their solar flare exposure ages which agree with the 'surface' exposure ages based on ^{26}Al and particle track studies³. We describe here, for the first time, using ^3He in samples from three different depths of lunar rocks, a method for deducing the absolute SCR (flare) proton fluxes in the past few million years which compare favourably with those derived from ^{26}Al and ^{53}Mn studies^{3,6} of other lunar samples.

When a large crater is formed on the surface of the Moon, deeply buried rocks are excavated to the surface. During such shock events, rocks could be either completely or partially degassed of their noble gas. The helium in a heavily shocked rock is considered to have completely degassed. As the earlier ^3He history in this rock was completely erased out at the time it was brought to the lunar surface by the shock event, its subsequent irradiation on the lunar surface results in equal exposure times to SCR and GCR. If $m(^3\text{He})$ is the measured

spallogenic ^3He content per g of the rock with a simple history, we have

$$m(^3\text{He}) = (P_g + P_s)T$$

where P_g and P_s are the GCR and SCR ^3He production rates respectively. Choosing a deep interior sample (R3) from a rock, such that $P_s = 0$, means that $m(^3\text{He}) = P_g T$. As P_g is known, T for R3 can be calculated. Samples close to the top surface (R1) will contain a mixture of SCR, GCR, SW (solar wind) components of ^3He and intermediate samples (R2) deeper than 3–4 mm from the top surface will contain only GCR and SCR produced ^3He . The ^3He in R1 sample is corrected for solar wind contribution using the measured ^4He content and a $^3\text{He}/^4\text{He}$ ratio of 3.6×10^{-4} as found in several bulk lunar soils⁷. The corrected ^3He in R1 represents the SCR + GCR production.

In the case of rocks with a simple history, the SCR proton flux is deduced as follows. Assuming (1) $T(\text{GCR}) = T(\text{SCR})$ and (2) the diffusion losses, if any, are same for both GCR and SCR produced ^3He , we can calculate $T(\text{GCR})$ from R3 which is equal to $m(^3\text{He})/P_g$ as shown above. T will decrease if there is any gas loss and irrespective of such losses, $m(^3\text{He})/T$ will give the true production rate. Using the calculated T , we compute the total SCR + GCR production rates, $(P_g + P_s) = m(^3\text{He})/T$ for other samples (R1 and R2) and plot these values against depth as shown in Fig. 1. Then, using assumed SCR fluxes and the known GCR flux, we calculate the combined production profiles for various depths in the rock and match the production profiles with the data points. The SCR parameters that fit best to the data points describe quantitatively the SCR flux to which this rock was exposed.

We measured He, Ne, Ar contents in six samples from known depths of two well documented rocks 61016 and 64435. Using the analytical methods described elsewhere² we could resolve quantitatively the SW, SCR and GCR components for the ^{21}Ne and ^{38}Ar in these samples. Using the production rates of Hohenberg *et al.*⁸, we calculated their SCR and GCR ages (Table 1) and the results obtained by both ^{21}Ne and ^{38}Ar methods show agreement within experimental errors. The measured ^{40}Ar contents are much less compared with the

Table 1 GCR and SCR ages of lunar rocks

Rock	^{21}Ne age (Myr)		^{38}Ar age (Myr)	
	SCR	GCR	SCR	GCR
61016*	1.3 ± 0.2	4.0 ± 0.3	1.7 ± 0.2	3.7 ± 0.3
64435	0.6 ± 0.1	1.2 ± 0.2	0.8 ± 0.2	1.4 ± 0.2

R1, R2 and R3 samples of 61016 are taken from 0.3–2 mm; 4–7 mm and 16–18 mm depths whereas another set of R1, R2 and R3 samples of 64435 are taken from 0.2–1.5 mm; 4.5–7.5 mm and 7.5–9.0 mm depths, respectively.

* See ref. 2.

expected ^{40}Ar from ^{40}K decay during 3,100–3,400 Myr after solidification⁹. We attribute the loss of ^{40}Ar (about 90%) to the shock event at the time of South Ray crater formation that excavated the rocks from lunar interior. During the same event more than 99% of the radiogenic ^4He was also lost. Thus the assumption that $T(\text{GCR}) = T(\text{SCR})$ can be applied to the rock samples from 61016 and 64435.

As pointed out earlier, the $T(\text{GCR})$ obtained from $m(^3\text{He})/P_g$ represents the apparent age of the rock and it will be lower than the true value if there are ^3He losses. Helium may sometimes be lost by diffusion and this source of gas loss is discussed below. To study the depth dependence of ^3He diffusion, we consider the R2 and R3 results of rock 64435. The observed ^3He contents are $(0.44 \pm 0.02) \times 10^{-8}$ and $(0.43 \pm 0.03) \times 10^{-8} \text{ cm}^3 \text{ STP g}^{-1}$ respectively and they do not show much variation as a function of depth. Further, a study of $^3\text{He}/^{21}\text{Ne}$ and $^3\text{He}/^{38}\text{Ar}$ ratios in these samples may throw light on the diffusion pattern of ^3He relative to other cosmogenic isotopes of Ne and Ar. The ratio of the $^3\text{He}(\text{SCR} + \text{GCR})/^{21}\text{Ne}(\text{SCR} + \text{GCR})$ in different samples of rocks 61016 and 64435 range from 3.5 to 4.0. The near-constancy of the observed $^3\text{He}/^{21}\text{Ne}$ ratios suggest that the depth-dependent diffusion effects are not significant in these samples. Similar observations are made by studying $^3\text{He}/^{38}\text{Ar}$ ratios in these samples from rocks 61016 and 64435. In this context, note that ^3He losses were observed in some feldspar-bearing lunar¹⁰ rocks and meteorites¹¹. It is not clear whether these losses are due to diffusion or to shock effects. For example, Heymann *et al.*¹¹ showed that in meteorites such as Shergotty and Zagami, even though feldspar-rich, the ^3He was almost quantitatively retained. In Sera de Mage also, no loss of ^3He in big feldspar grains was observed. This suggests that shock effects have an important role in affecting ^3He losses in some of these samples. However, even if there are minor diffusion losses, this method holds good as long as the losses are not depth-dependent.

As we do not have the deepest samples where P_s is almost zero, we adopt the following procedure to deduce the SCR proton fluxes from our studies. We calculate the theoretical ^3He production profile as a function of depth for combined SCR + GCR production per Myr, using the Reedy–Arnold model¹² and a SCR proton flux of $J(E > 10 \text{ MeV}) = 70 \text{ p cm}^{-2} \text{ s}^{-1}$ and $R_0 = 100 \text{ MV}$ (ref. 6) and a GCR flux of $1.7 \text{ p cm}^{-2} \text{ s}^{-1}$. The production cross-sections are taken from ref. 13. Figure 1a represents this theoretical profile. Using the SCR + GCR combined production rate for the 61016–R3 sample at a depth of 5.1 g cm^{-2} , we calculate $T(\text{GCR})$ for 61016 to be $(1.6 \pm 0.1) \text{ Myr}$. This agrees with the SCR ages based on ^{21}Ne and ^{38}Ar (see Table 1) and the surface age of 1.5 Myr based on particle track studies³. Then we use this $T(\text{GCR})$ to find the $m(^3\text{He})/T$ for the R1 sample of 61016. The observed R3 value is normalized so as to fall on curves a or b in Fig. 1 and the other data points are scaled accordingly. These values are plotted in Fig. 1. Similarly in case of 64435–R3 sample, using the (SCR + GCR) ^3He production rate at the corresponding depth, we calculate the $T(\text{GCR})$ for rock 64435 which turns out to be $(0.55 \pm 0.10) \text{ Myr}$. This agrees with the SCR ages based on ^{21}Ne and ^{38}Ar within the error limits (see Table 1). We plot the $m(^3\text{He})/T$ values for the R2 and R1 samples in Fig. 1 as a function of depth as described above. Within the error limits, all the points of rocks 61016 and 64435 seem to fit curve Fig. 1a.

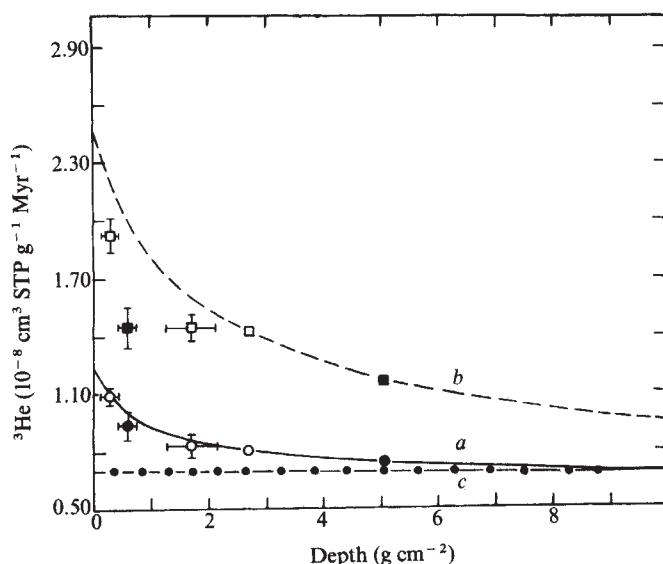


Fig. 1 Production profiles of (SCR + GCR) ^3He as a function of depth in lunar rocks 64435 ($\square$, $\circ$) and 61016 ($\blacksquare$, $\bullet$). The deepest points (R3) of 64435 (at 2.7 g cm^{-2}) and of 61016 (at 5.1 g cm^{-2}) are normalized to fall on curves a or b. The other R1 and R2 points are scaled accordingly. The normalizing points are shown without error bars. The solid curve a is generated from the SCR proton flux given by Kohi *et al.*⁶ and the dashed curve b is generated from the SCR flux given by Bhandari *et al.*³. Curve c represents the depth profile¹² for GCR production of ^3He using the particle flux mentioned for proton energies above 1 GeV. ^3He in R2 sample of 61016 could not be measured due to power failure.

We can then check how these data points, using the same procedure, fit to a different SCR proton flux of J ($E > 10$ MeV) = $125 \text{ p cm}^{-2} \text{ s}^{-1}$ and $R_0 = 150$ MV, deduced by Bhandari *et al.*³ for the same rocks based on ^{26}Al studies. The combined (SCR + GCR) ^3He production profile calculated for this SCR flux is shown by Fig. 1b. The T (SCR) calculated for this flux for 61016 turns out to be 1.0 Myr and for 64435, 0.3 Myr which are smaller than the surface exposure ages of 1.5 Myr and 0.5 Myr respectively, deduced by track studies³. Using the normalization procedure, for R3 points of rocks 61016 and 64435, we calculate the other R1 and R2 points with the new (SCR + GCR) ^3He production rates and plot the points as a function of depth in Fig. 1.

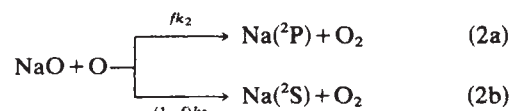
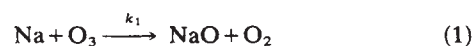
The fit of the data points to curve *b* is clearly less satisfactory than to curve *a*. This shows that the ^3He production profile generated using a SCR proton flux of J ($E > 10$ MeV) = $70 \text{ p cm}^{-2} \text{ s}^{-1}$ and $R_0 = 100$ MV (ref. 6) seems to fit the data points better. The erosion, solid angle and so on, corrections for these samples are expected to be minor. Further experiments are in progress using well-documented samples to understand this discrepancy in the fluxes as well as to find the suitability of this method to determine the solar flare exposure ages of lunar samples in time scales > 10 Myr where the radioactivity methods using ^{26}Al and ^{53}Mn are inapplicable because of saturation effects.

We thank Professors D. Lal and N. Bhandari for valuable discussions during this work which is carried out under Lunar Science Proposal Nos. SL-9191 and SL-9095, and J. T. Padia for assistance.

Received 19 March; accepted 26 June 1980.

1. Bhai, N. B., Gopalan, K., Goswami, J. N., Rao, M. N. & Venkatesan, T. R. *Proc. 9th Lunar planet. Sci. Conf.* 1629–1645 (1978).
2. Rao, M. N., Venkatesan, T. R., Goswami, J. N., Nautiyal, C. M. & Padia, J. T. *Proc. 10th Lunar planet. Sci. Conf.* 1547–1564 (1979).
3. Bhandari, N., Bhattacharya, S. K. & Padia, J. T. *Proc. 7th Lunar Sci. Conf.* 513–523 (1976).
4. Arrhenius, G. *et al. Proc. 2nd Lunar Sci. Conf.* 2583–2589 (1971).
5. Hoyt, H. P., Jr., Walker, R. M. & Zimmerman, D. W. *Proc. 4th Lunar Sci. Conf.* 2489–2502 (1973).
6. Kohl, C. P., Murrell, M. T., Russ III, G. P. & Arnold, J. R. *Proc. 9th Lunar planet. Sci. Conf.* 2299–2310 (1978).
7. Eberhardt, P. *et al. Proc. 3rd Lunar Sci. Conf.* 1821–1856 (1972).
8. Hohenberg, C. M., Marti, K., Podosek, F. A., Reedy, R. C. & Shirck, J. R. *Proc. 9th Lunar planet. Sci. Conf.* 2311–2344 (1978).
9. Stettler, A., Eberhardt, P., Geiss, J., Grogler, N. & Maurer, P. *Proc. 4th Lunar Sci. Conf.* 1865–1888 (1973).
10. Lugmair, G. W., Marti, K., Kurtz, J. P. & Scheinin, N. B. *Proc. 7th Lunar Sci. Conf.* 2009–2033 (1976).
11. Heymann, D., Mazor, E. & Anders, E. *Geochim. cosmochim. Acta* **32**, 1241–1268 (1968).
12. Reedy, R. C. & Arnold, J. R. *J. geophys. Res.* **77**, 537–555 (1972).
13. Walton, J. R., Heymann, D., Yaniv, A., Edgerley, D. & Rowe, M. W. *J. geophys. Res.* **81**, 5689–5699 (1976).

In 1939 Chapman¹ proposed the oxidation–reduction cycle:



Thermochemical studies² have removed doubts over whether process (2a) is exothermic.

The case against the possibility of invoking another chemi-luminescent cycle may be seen to be virtually conclusive: thus rocket measurements³ show that the emission comes from near the 90-km level and lidar measurements⁴ show that the number density $n(\text{Na})$ in this region is around $4 \times 10^3 \text{ cm}^{-3}$, so that even with a rate coefficient as fast as $2.5 \times 10^{-10} \text{ cm}^3 \text{ s}^{-1}$ and a 10-km thick layer, the typical nightglow intensity⁵ of 100 R would necessitate the number density of the reactant being some $1 \times 10^8 \text{ cm}^{-3}$; ozone is almost certainly the only atmospheric constituent which readily reacts with or excites atomic sodium and which may have such a high nocturnal abundance near 90 km; and, if it is accepted that process (1) must be involved, then so almost certainly must process (2) because of the lack of another reducing process which meets the double requirement of being at least as effective in the atmospheric region concerned and of being sufficiently exothermic. Furthermore, interferometric measurements on D-line profiles⁶ reveal that the $\text{Na}(\text{P})$ atoms have a kinetic energy of 0.046 eV which suggests that they arise from a chemical process like process (2a) rather than from dissociative recombination involving NaO^+ or NaO_2^+ produced from incoming Na^+ ions as has been conjectured⁷.

An understanding of the nature of processes (1) and (2) came when Kolb and Elgin⁸ suggested that both are of the electron-jump type. From the ionization potentials of Na and NaO and the electron affinities of O_3 and O they calculated the crossing distances r_x for the $\{\text{Na}-\text{O}_3\} : \{\text{Na}^+-\text{O}_3^-\}$ and $\{\text{NaO}-\text{O}\} : \{\text{NaO}^+-\text{O}^-\}$ covalent–ionic potential pairs. Assuming the cross-sections to be πr_x^2 , they deduced the rate coefficients. The values of r_x are 4.5 Å and 2.3 Å respectively (the latter being smaller than originally⁸ obtained due to the use of what should be a better⁹ estimate of the ionization potential of NaO); the corresponding rate coefficients at 200 K are

$$k_1 = 3.4 \times 10^{-10} \text{ cm}^3 \text{ s}^{-1}, \quad k_2 = 1.0 \times 10^{-10} \text{ cm}^3 \text{ s}^{-1} \quad (3)$$

Note, however, that processes (1) and (2) are not of the class typified by



to which the electron-jump model is commonly applied^{10,11}. A key feature of process (4) is that the I_2 formed in a vertical transition at the crossing is so highly excited vibrationally that it dissociates readily leaving the I^- and K^+ ions bound together by their Coulomb attraction and leaving the I atom free. In the case of processes (1) and (2), there is little vibrational excitation in a vertical transition and, moreover, the product molecules are not produced directly from oppositely charged fragments. The mechanism is that, while the reactants are on the ionic potential surface, they make a hard collision which excites internal energy and thereby makes an exothermic channel more likely to be followed than reversal through the entrance channel. For process (1), the fission of the collision complex occurs along $\{\text{NaO}^+-\text{O}_2\} : \{\text{NaO}-\text{O}_2\}$ and for process (2), along $\{\text{Na}^+-\text{O}_2\} : \{\text{Na}-\text{O}_2\}$, where the $\{\text{NaO}^+-\text{O}_2\}$ and $\{\text{Na}^+-\text{O}_2\}$ potential surfaces are for the same electronic states as the $\{\text{Na}^+-\text{O}_3\}$ and $\{\text{NaO}^+-\text{O}\}$ potential surfaces, respectively.

Oxidation and reduction proceeds rapidly enough to maintain chemical equilibrium between Na and NaO and it is convenient to express the D-doublet photon emission rate as

$$E(5,893) = fk_1 n(\text{Na})n(\text{O}_3) \quad (5)$$

Excitation of the Na D-doublet of the nightglow

D. R. Bates & P. C. Ojha

Department of Applied Mathematics and Theoretical Physics, Queen's University Belfast, Belfast BT7 1NN, UK

We infer from the absolute intensity of the nocturnal D-doublet emission, the abundance of free Na and other relevant information that there is very little possibility of an alternative to Chapman's theory¹ that the excited Na atoms involved are released by $\text{NaO}-\text{O}$ collisions. The nature of the collision process is discussed here. We point out that a remarkably high branching ratio for the reaction path leading to excited (rather than normal) sodium is necessary. A high branching ratio is explained by considering the symmetry properties of the correlated electronic states.

Table 1 Symmetry properties of the Na-O correlated molecular states*

Row	Fragments	Correlated molecular states*
(a)	$\text{NaO}(\text{X}^2\Pi) + \text{O}^3(\text{P})$	$^2\text{A}'(3), ^4\text{A}'(3), ^2\text{A}''(3), ^4\text{A}''(3)$
(b)	$\text{NaO}^+(\text{X}^3\Sigma^-) + \text{O}^-(^2\text{P})$	$^2\text{A}'(1), ^4\text{A}'(1), ^2\text{A}''(2), ^4\text{A}''(2)$
(c)	$\text{Na}^+(^1\text{S}) + \text{O}_2^-(\text{X}^2\Pi_g)$	$^2\text{A}'(1), ^2\text{A}''(1)$
(d)†	$\text{Na}^+(^1\text{S}) + \text{O}_2^-(^4\Sigma_g^+, ^4\Delta_g)$	$^4\text{A}'(2), ^4\text{A}''(1)$
(e)	$\text{Na}(^2\text{S}) + \text{O}_2(\text{X}^3\Sigma_g^-)$	$^2\text{A}''(1), ^4\text{A}''(1)$
(f)	$\text{Na}(^2\text{S}) + \text{O}_2(a^1\Delta_g)$	$^2\text{A}'(1), ^2\text{A}''(1)$
(g)	$\text{Na}(^2\text{S}) + \text{O}_2(b^1\Sigma_g^+)$	$^2\text{A}'(1)$
(h)	$\text{Na}(^2\text{P}) + \text{O}_2(\text{X}^3\Sigma_g^-)$	$^2\text{A}'(1), ^4\text{A}'(1), ^2\text{A}''(2), ^4\text{A}''(2)$

* Following the standard notation¹⁸, A' indicates a state whose eigenfunction is symmetrical with respect to the plane and A'' one whose eigenfunction is anti-symmetrical while the numeral in brackets is the number of such states.

† Only the two lowest¹⁸ suitable states of O_2^- are displayed.

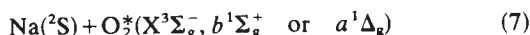
the number densities on the right being better determined than the other number densities concerned. Kolb and Elgin⁸ did not discuss the factors which control the branching ratio f but commented that a value between 10^{-1} and 10^{-2} would apparently be required to account for the nightglow. Recent data requires that f be not greatly below unity.

Stellar occultation measurements by Reigler and others^{12,13} show that near the Equator $n(\text{O}_3)$ is some $1 \times 10^8 \text{ cm}^{-3}$ or less between 85 and 90 km rising to $\sim 2 \times 10^8 \text{ cm}^{-3}$ at 95 km. Real variations occur and considerable uncertainty remains. Measurements¹⁴ (31°N, February) on the 110.8 GHz O_3 rotational line give $5 \times 10^8 \text{ cm}^{-3}$ at 80 km falling to about $1 \times 10^8 \text{ cm}^{-3}$ at 85 km; indirect deduction¹⁵ from satellite data (20°N, June) on the intensity of the (9, 3) hydroxyl band give $6 \times 10^8 \text{ cm}^{-3}$ between 80 and 85 km and $3 \times 10^8 \text{ cm}^{-3}$ at 90 km. Atmospheric model calculations¹³ give a nocturnal distribution with a broad maximum of around $2 \times 10^8 \text{ cm}^{-3}$ in the 85–95 km region. Simultaneous lidar measurements on $n(\text{Na})$ and photometer measurements on the intensity $I(5,893)$ of the D-doublet have been carried out at São José dos Campos (23°S)¹⁶. The inference was drawn that the peak photon emission rate was at 88 km, where $n(\text{Na})$ was $3 \times 10^3 \text{ cm}^{-3}$, and had the value $1 \times 10^2 \text{ cm}^{-3} \text{ s}^{-1}$. To reproduce this, it may be seen from process (5) that, with k_1 as in process (3), f would have to be about 0.3 even if $n(\text{O}_3)$ were as great as $3 \times 10^8 \text{ cm}^{-3}$. Later lidar and photometer measurements¹⁷ lead to compatible values. We need to examine whether f for process (2) can indeed be as high as required for the nightglow.

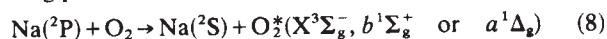
Parting along $\{\text{Na}^+\text{O}_2\} : \{\text{Na-O}_2\}$ may yield either



or



the asterisk indicating possible vibrational excitation. The crossings giving rise to reaction products (6) are at a greater separation than the numerous crossings giving rise to reaction products (7) and these latter are not readily traversed as is evident from the large¹⁸ cross-section ($34\text{--}38 \text{ \AA}^2$) for the quenching process



Furthermore, compared with the channels yielding products (6), the many relatively exothermic channels yielding products (7) are very strongly favoured statistically. Thus, it might seem that f must be minute.

However, consider the symmetry properties of the correlated molecular states concerned. Referring to Table 1, species $^4\text{A}'$ appears in reactant row (a), the product row (h) and the intermediate ionic rows (b) and (d) the last of which involves excited O_2^- , but does not appear in the product rows (e) to (g). Because the species does not change during an encounter, it follows that the simple electron-jump model gives the statistical weight ratio of 1/3 for f the branching ratio. Although the

closeness of the fragments at the crossings may affect the quantitative predictions of the model regarding k_2 (which is not of critical importance) and f , there is certainly no reason to doubt that the branching ratio is as high as Chapman's theory demands.

D.R.B. thanks Professors T. M. Donahue and M. Nicolet for correspondence on the nocturnal ozone abundance.

Received 2 May; accepted 26 June 1980.

1. Chapman, S. *Astrophys. J.* **90**, 309–316 (1939).
2. Hildenbrand, D. L. & Murad, E. *J. chem. Phys.* **53**, 3403–3408 (1970).
3. Greer, R. G. H. & Best, G. T. *Planet. Space Sci.* **15**, 1857–1881 (1967).
4. Sandford, M. C. W. & Gibson, A. J. *J. atmos. terr. Phys.* **32**, 1423–1430 (1970).
5. Hunten, D. M. in *The Radiating Atmosphere* (ed. McCormac, B. M.) 3–16 (Reidel, Dordrecht, 1971).
6. Sipler, D. P. & Biondi, M. D. *Planet. Space Sci.* **26**, 65–73 (1978).
7. Sechrist, C. F. *EOS* **57**, 972 (1976).
8. Kolb, C. E. & Elgin, J. B. *Nature* **263**, 488–489 (1976).
9. O'Hare, P. A. G. & Wahl, A. C. *J. chem. Phys.* **56**, 4516–4525 (1971).
10. Herschbach, D. R. *Adv. Chem. Phys.* **10**, 319–392 (1966).
11. Gislason, E. A. in *Alkali Halide Vapors* (eds Davidovits, R. & McFadden, C. L.) 415–440 (Academic, New York, 1979).
12. Reigler, G. R., Drake, J. F., Liu, S. C. & Cicerone, R. J. *J. geophys. Res.* **81**, 4997–5001 (1976).
13. Reigler, G. R. *et al. Geophys. Res. Lett.* **4**, 145–148 (1977).
14. Penfield, H., Litvak, M. M., Gottlieb, C. A. & Lilley, A. E. *J. geophys. Res.* **81**, 6115–6120 (1976).
15. Frederick, J. E., Rusch, D. W. & Liu, S. C. *J. geophys. Res.* **83**, 2441–2445 (1978).
16. Clemesha, B. R., Kirchhoff, V. W. J. H. & Simonich, D. M. *J. geophys. Res.* **83**, 2499–2503 (1978).
17. Kirchhoff, V. W. J. H., Clemesha, B. R. & Simonich, D. M. *J. geophys. Res.* **84**, 1323–1327 (1979).
18. Grice, R., *Adv. Chem. Phys.* **30**, 247–312 (1975).
19. Krauss, M., Neumann, D., Wahl, A. C., Das, G. & Zemke, W. *Phys. Rev. A* **7**, 69–77 (1973).
20. Herzberg, G. *Electronic Spectra of Polyatomic Molecules*, 10 (Van Nostrand, Princeton, 1966).

Physicochemical speciation of lead in drinking water

Roy M. Harrison & Duncan P. H. Laxen

Department of Environmental Sciences, University of Lancaster, Lancaster LA1 4YQ

Recent studies have highlighted the importance of drinking water as a route of human exposure to lead^{1,2}. Whilst there are ample data on lead concentrations in drinking water³, little is known of its physical and chemical forms (physicochemical speciation)⁴. Such information is important as the speciation of ingested lead influences the efficiency of absorption from the gastrointestinal tract⁵. Knowledge of speciation should also provide a fuller understanding of the factors controlling the solubility of lead in potable waters⁶ and hence assist in devising the most cost-effective means of plumbosolvency control. We have determined experimentally the speciation of lead in three different tapwaters and report here diverse forms of dissolved and particle-associated lead, dependent primarily on the chemical matrix of the raw water.

Based on calculations of equilibrium solubilities, the Water Research Centre has predicted that basic lead carbonate, $2\text{PbCO}_3 \cdot \text{Pb}(\text{OH})_2$ is the most stable insoluble lead compound to be formed in the conditions pertaining to most potable waters in Britain⁶. We have analysed the off-white surface deposit from inside a lead water pipe using X-ray powder diffraction and confirm this to be basic lead carbonate. However, as the Water Research Centre indicated⁶, higher concentrations of lead can occur in potable waters than would be predicted by equilibrium of free lead ions with basic lead carbonate. The formation of ion pairs, or organic complexes of lead, or the presence of lead in insoluble particulate form in the water were all suggested as possible means by which these enhanced concentrations might occur, although no experimental measurements were made.

A range of techniques suitable for metal speciation exist, all of which essentially classify the species according to some operationally defined association⁷. We have recently developed a

Table 1 Analysis of tapwater samples

Sample	Alkalinity ($\mu\text{g ml}^{-1}$ as CaCO_3)	pH	Organic carbon ($\mu\text{g ml}^{-1}$)	Specific conductance (μS)
Lancaster University	21	7.5	9.4	130
Bentham	20	6.8	1.6	92
Glasgow	4	6.7	8.0	42

speciation scheme for freshwaters, which combines several established techniques⁸. It is based principally on a size fractionation of the metal species. Briefly, samples were collected in polythene bottles and filtered as soon as possible (in parallel, not in series) through Nuclepore filters under pressure in Amicon stirred cells. In two cases filtration was started within 2 h of collection, whilst with the Glasgow sample filtration was unavoidably delayed for 24 h. Nuclepore filters were selected as they have inherently better size discrimination properties than membrane filters and suffer lesser contamination and adsorption problems. Both filter and filtration apparatus were pretreated with a $\text{Ca}(\text{NO}_3)_2$ solution to minimize adsorption losses⁹. They were rinsed with deionized double-distilled water to remove excess $\text{Ca}(\text{NO}_3)_2$ and then with a portion of sample before filtration. Recovery tests with pre-filtered samples of tapwater and riverwater, as well as with a synthetic river water, using filters of 0.015–12 μm pore size showed recovery to be independent of filter pore size. Mean recoveries were as follows: Pb, 86%; Cu, 94%; Fe, 101%; Mn, 100%. A fuller discussion of the significance of these adsorption tests is given elsewhere⁸. Note, however, that if, as is likely, adsorption is independent of filter pore size then adsorption losses will only affect the quantification of the smallest size fraction, as the others are calculated by difference and are hence independent of adsorption losses. Contamination during filtration was very low: Pb, $<0.1 \text{ ng ml}^{-1}$; Cu, $<0.1 \text{ ng ml}^{-1}$; Fe, $<0.5 \text{ ng ml}^{-1}$; Mn, $<0.05 \text{ ng ml}^{-1}$.

The total metal concentrations in each filtrate were determined by flameless atomic absorption for Fe and Mn, and by anodic stripping voltammetry (ASV) for Pb and Cu following UV-ray acid digestion. A standard additions calibration was applied in all cases. As a measure of the ASV-labile concentration at the natural pH of the sample, ASV analyses were performed on the 1- μm filtrate at a hanging mercury drop electrode, using CO_2 in nitrogen as a buffer. A titration procedure was used to indicate the extent and strength of any complexation and the slope of the titration curve after saturation of the complexation capacity was used to quantify the labile response.

Labile metals were also defined by their uptake on Chelex 100 exchange resin. A 48-h batch technique was employed, using 2 ml of resin in the calcium form per 60 ml of sample in a polythene bottle. This provided a measure of metal complexes with slower dissociation kinetics. Finally, UV-ray irradiation (1 kW Hanovia lamp) was used to destroy organic matter and so release metals previously associated with the organics in a non-ASV-labile form. The increase in labile metal following UV-ray irradiation is a measure of the metals strongly bound with organic matter⁷. This procedure has, however, proved to be inapplicable for most freshwaters, there being a loss rather than gain of ASV-labile metal, due probably to uptake onto hydrous ferric oxide precipitated during the irradiation⁸. It still, nevertheless, provides qualitative information on the nature of the ASV-labile metal.

Particular care has been exercised in the evaluation of the ASV-labile results, due to recently recognized limitations of the method^{10,11}. However, interpretation of the data is strengthened by a consistency between the results derived from each of the techniques. Finally, note that in common with all speciation studies, the results might to some unknown extent reflect artefacts of the various handling operations, starting with the sampling itself.

Three tapwater samples were analysed. These included two morning first draw waters, from Glasgow (Scotland) and Bentham (North Yorkshire). The former came from a house with lead-lined tank and lead piping, the latter from a house with copper piping internally, but a lead service pipe. Both waters were of rather low alkalinity and pH (Table 1) and could be described as aggressive in respect of plumbosolvency. The other sample was from the University of Lancaster (copper piping with high-lead soldered capillary joints) and was taken after a thorough flushing of the pipes.

The results are presented in Table 2. Copper has been included, as its speciation behaviour usually differs from that of lead, the latter forming predominantly inorganic species, the former organic¹². As a substantial proportion of the lead appeared in the colloidal size fractions, the elements iron and manganese were also analysed, as these elements can exist in waters as colloidal hydrous oxide particles or coatings, as well as in association with organic macromolecules¹³. The colloidal hydrous oxides are effective at adsorbing metals from waters and show a particularly strong affinity for lead¹⁴. The University tapwater sample had a total lead concentration of only 3.3 ng ml^{-1} , the major proportion of which, 76%, was filterable ($<0.015 \mu\text{m}$). This filterable lead was however not ASV-labile, and hence cannot exist as Pb^{2+} . The water supply is a mixture of moorland surface water and River Lune water and has a substantial organic content (Table 1). Titration of the tapwater sample with lead showed further ability to complex lead as indicated by the ASV signal until the complexation capacity had been exceeded. This complexation capacity was lost following UV-ray irradiation of the sample, which destroys organic matter. The major portion of the lead is, therefore, probably present as non-labile low molecular weight organic complexes, although some may also exist as non-labile inorganic species.

In the Bentham tapwater, which is derived from spring waters and is of low organic content, 53% of the lead and 33% of the copper were associated with colloidal particles of 0.015–1.0 μm . A significant fraction of the lead (19%) was in the particulate size range 1–12 μm . This may have been due to basic lead carbonate which precipitates as rather large particles (unpublished results). The size distribution of the lead in the colloidal size range parallels that of iron rather closely. The mole ratio Pb/Fe is ~ 0.2 in all three fractions between 0.015 and 1.0 μm . This is consistent with the lead occurring in association with hydrous oxides of iron, either by ion-exchange, by specific adsorption or by co-precipitation. The Chelex-labile copper corresponds closely to the ASV-labile metal, whilst for lead more metal is Chelex-labile. This additional labile metal must, furthermore, be due in some way to release from the colloidal metal fraction. It is either a result of weak lead complexation with organic macromolecules or a slow release from the hydrous oxides due to the disturbed equilibrium in the presence of the Chelex resin. There was no apparent complexation capacity for additional lead during titration of the sample, and UV-ray irradiation did not cause a change in the peak potential of the ASV lead signal, adding weight to the suggestion that inorganic species of lead predominate in this water.

The Glasgow sample, abstracted originally from Loch Katrine, contained the highest concentration of total lead, 372 ng ml^{-1} , and the highest percentage of ASV-labile lead, 63%. Again a high proportion (71%) of lead was associated with the colloidal size fractions, although in contrast to the Bentham water most of this colloidal lead (75%) was in the small size range 0.015–0.08 μm . Note that a significant proportion of this colloidal lead must contribute to the ASV and Chelex-labile lead. During titration of the sample with additional lead a separate lead signal at a more anodic potential developed after $\sim 800 \text{ ng ml}^{-1}$ of lead had been added. This is indicative of complexation of the ASV-labile metal in the original sample¹⁵. After UV-ray irradiation of the sample the original ASV signal disappeared to be replaced by a more anodic peak. It seems that a large proportion of the lead in this sample (63%) existed as ASV-labile organic complexes, a major portion being of high

molecular weight (0.015 μm is equivalent to a molecular weight of $\sim 300,000$). In the case of copper in the raw sample, ASV-labile metal could not be distinguished. Copper forms stronger complexes with organics than lead and its complexation is often non-ASV-labile¹², as seems to be the case with this sample.

In the three waters examined there is apparently little or no free ionic lead. A significant portion of the lead exists in the colloidal size fractions, associated with hydrous iron oxides or organic macromolecules depending on the composition of the water. There is some particulate lead in the 1–12 μm size range, which may be metal precipitate, or may be associated with clay minerals, but little or no lead in larger particles, $>12 \mu\text{m}$. This would suggest that the lead in the samples examined is not a result of particles flaking off the basic lead carbonate lining of the lead water pipes.

Plumbosolvency is clearly not a simple equilibrium between solid basic lead carbonate and dissolved Pb^{2+} ions and ion pairs. It depends rather on the adsorption and complexation capacity of the hydrous oxides and organic complexing agents such as humic acids. Techniques of plumbosolvency control depend on an increase in the sample pH and hardness/alkalinity^{6,16}. The lead concentration is, however, not always reduced to the extent predicted by theoretical calculations of equilibrium solubility^{6,16} and the complex lead speciation in real tapwaters, as indicated by our results, may account for this result. The reduction in

plumbosolvency achieved by pH and hardness/alkalinity control may be a function of a decrease in the rate of lead dissolution and/or the effect of competing cations such as Ca^{2+} for adsorption, ion-exchange and complexation sites, in addition to the anticipated decrease in equilibrium solubility.

One option that has been considered for reduction of lead levels in drinking water is the use of chelating resin at the consumer's tap⁶. Our results demonstrate that this measure is likely to have a limited effectiveness, as a substantial fraction of the lead will probably be in a non-ion-exchangeable form.

The health implications of our findings are uncertain. However, at pH 1–2 in the acidified samples, all lead is ASV-labile (Table 2) and is probably present as free metal ions. Hence at the pH of the stomach, tapwater lead will also take this form and hence may well be more available for absorption than the lead bound in foodstuffs. This may explain the high effectiveness of removal of lead piping in reducing blood lead levels in the study previously cited².

We thank Dr M. R. Moore and Mr A. Britten for assistance in providing the Glasgow tapwater sample.

Received 6 May; accepted 2 July 1980.

Table 2 Results of speciation study

a, Lancaster University tapwater, well flushed

Size fraction (μm)	Metal concentration (ng ml ⁻¹ (%))			
	Pb	Cu	Fe	Mn
>12	—	2 (8)	17 (9)	1.4 (16)
1.0–12	—	—	—	—
0.4–1.0	—	—	—	—
0.08–0.4	0.8 (24)	3 (12)	173 (87)	2.7 (30)
0.015–0.08				
<0.015	2.5 (76)	19 (79)	8 (4)	4.8 (54)
Total	3.3	24	198	8.9
ASV-labile*	<0.2 (<6)	1 (4)	NA	NA

b, Bentham tapwater, first flush

Size fraction (μm)	Metal concentration (ng ml ⁻¹ (%))			
	Pb	Cu	Fe	Mn
>12	—	—	51 (30)	11.1 (53)
1.0–12	16 (19)	12 (3)	51 (30)	—
0.4–1.0	14 (17)	40 (11)	16 (10)	—
0.08–0.4	22 (26)	42 (12)	34 (20)	0.8 (4)
0.015–0.08	8 (10)	36 (10)	10 (6)	1.8 (9)
<0.015	24 (29)	228 (64)	6 (4)	7.1 (34)
Total	84	358	168	20.8
ASV-labile*	20 (24)	231 (65)	NA	NA
Chelex-labile*	39 (46)	224 (63)	20 (12)	7.3 (36)
Acid ASV-labile†	87 (104)	360 (101)	NA	NA

c, Glasgow tapwater, first flush

Size fraction (μm)	Metal concentration (ng ml ⁻¹ (%))			
	Pb	Cu	Fe	Mn
>12	3 (1)	0.1 (<1)	1.8 (3)	0.3 (7)
1.0–12	49 (13)	1.7 (8)	20.1 (35)	2.1 (47)
0.4–1.0	7 (2)	0.9 (4)	2.4 (4)	0.2 (4)
0.08–0.4	59 (16)	1.9 (8)	6.1 (11)	0.4 (9)
0.015–0.08	198 (53)	9.5 (42)	18.0 (31)	0.3 (7)
<0.015	56 (15)	8.3 (37)	9.0 (16)	1.2 (27)
Total	372	22.4	57.4	4.5
ASV-labile*	236 (63)	NA	NA	NA
Chelex-labile*	207 (56)	9.8 (44)	5.9 (10)	1.1 (24)
Acid ASV-labile†	382 (103)	22.4 (100)	NA	NA

* 1- μm filtrate.

† Unfiltered sample 0.5% HNO_3 .

NA, Not available.

1. Department of Health and Social Security *Lead and Health* (HMSO, London, 1980).
2. Thomas, H. F., Elwood, P. C., Welsby, E. & St Leger, A. S. *Nature* **282**, 712 (1979).
3. Department of the Environment *Pollution Paper No. 12* (HMSO, London, 1977).
4. *Health Hazards from Drinking-Water* (WHO, Copenhagen, 1978).
5. Chamberlain, A. C. *et al.* *UKAEA Rep. AERE-R9198* (HMSO, London, 1978).
6. Water Research Centre *Tech. Rep. TR43* (Water Research Centre, Medmenham, 1977).
7. Florence T. M. & Batley, G. E. *Talanta* **24**, 151 (1977).
8. Laxen, D. P. H. & Harrison, R. M. (in preparation).
9. Nürnberg, H. W., Valenta, P., Mart, L., Raspor, B. & Sipus, L. Z. *anal. Chem.* **282**, 357 (1976).
10. Brezonik, P. L., Brauner, P. A. & Stumm, W. *Water Res.* **10**, 605 (1976).
11. Shuman, M. S. & Michael, L. C. *Anal. Chem.* **50**, 2104 (1978).
12. Ernst, R., Allen, H. E. & Mancy, K. H. *Wat. Res.* **9**, 969 (1975).
13. Mill, A. J. B. *Envir. Techn. Lett.* **1**, 97 (1980).
14. Bacso, J., Szalay, A. & Kiss, K. *Sci. Tot. Envir.* **9**, 271 (1978).
15. O'Shea, T. & Mancy, K. H. *Anal. Chem.* **48**, 1603 (1976).
16. Akers, C. J. & Fellows, R. *Envir. Hlth* **87**, 148 (1979).

Measurements of CHFCl_2 in background tropospheric air

S. A. Penkett*, N. J. D. Prosser*, R. A. Rasmussen† & M. A. K. Khalil†

* Environmental and Medical Sciences Division, AERE Harwell, Oxon OX11 0RA, UK

† Oregon Graduate Center, Beaverton, Oregon.

The likely impact of fluorocarbons on the Earth's ozone layer has recently been reviewed in two major reports^{1,2} which, although showing difference of emphasis, both concluded that uncontrolled emissions of F11 (CFCl_3) and F12 (CF_2Cl_2) would lead to a substantial reduction in the level of ozone in the Earth's ozone shield (13–16%). The effect of F11 and F12 on ozone would be considerably reduced if they were removed on a large scale by some mechanism which operated in the troposphere. Such a mechanism has been proposed for F11, which involves conversion to F21 (CHFCl_2) on the surface of aerosols suspended in the atmosphere³. Crescentini and Bruner⁴ have recently reported measurements of the F21 concentration that were of sufficient magnitude to suggest that tropospheric removal of F11 was occurring on a scale greater than stratospheric removal. We report here data which refute this suggestion by showing that the level of F21 in the clean background troposphere is virtually unmeasurable.

Several attempts have been made to measure the concentration of F21 in the atmosphere. It was first detected by Rasmussen⁵ using a gas chromatograph/mass spectrometer (GC/MS) combination, and he quoted a value of ~ 10 p.p.t.v. (1 p.p.t.v. = 1 part in 10^{12} by volume). Singh *et al.*⁶ reported F21

background concentrations of 14.9 ± 4.9 p.p.t.v. as preliminary data for a clean site in Yosemite National Park, California. Subsequently, their more extensive global measurements⁷ in the Northern and Southern Hemispheres concluded a tropospheric average of 5 ± 3 p.p.t.v. However, they felt that the lack of a significant interhemispheric gradient for an anthropogenic compound with such a short atmospheric lifetime implied that it was probably an artefact of analysis, although they were unable to determine its source. Measurements in their studies were made using a gas chromatograph equipped with an electron capture detector (GC/ECD). F21 is a difficult species to measure and measurements made with the electron capture detector are not the most reliable as the detector is nonspecific. Glasgow *et al.*⁸ reviewed the artefact problems associated with these early F21 measurements. Crescentini and Bruner⁴ have shown that a species is present in air which has a mass spectrum identical to that of F21, and they estimate its concentration to be as high as 20 p.p.t.v. in air collected in central Italy. If this were the case for the well-mixed troposphere, then decomposition of F11 to F21 could represent a very substantial sink for F11. The present measurements using a GC/MS combination suggest that the level of F21 in the background troposphere is probably well below 1 p.p.t.

A Hewlett-Packard 5710 gas chromatograph and a VG micromass 16F mass spectrometer combination were used to make the measurements on concentrated air samples. Two concentration techniques were used, both based on collection of liquid air and subsequent removal of the major permanent gases at liquid nitrogen temperature. In one method the oxygen and nitrogen were removed by a stream of pure helium and carbonyl sulphide was used as an internal standard to estimate the effective volume injected (see ref. 9 where xenon was used for a similar purpose). The other method relied on carefully trapping a known volume of liquid air (~ 3 ml), then pumping away the oxygen/nitrogen excess. Alternatively, a small volume of gaseous air was injected into the cryotrap. Tests with a standard showed that a 95% recovery of material could be achieved by this method. The concentrated samples were then injected into a 3-m OV101 column held at 30°C , using helium as a carrier gas. COS was measured using mass 60 and F21 was measured using masses 67, 69 and 102. Because of the extremely low concentrations of F21 that were present, this was generally done using the single ion mode. Attempts to measure all three ions simultaneously resulted in rather poor signal-to-noise ratios. A typical chromatogram that resulted for mass 67 in the single ion mode is shown in Fig. 1. Chromatograms for masses 69 and 102 showed similar peak separation and signal-to-noise ratios. The peak ascribed to F21 in Fig. 1 had an identical retention time as an F21 standard, also a linear response to F21 concentrations between 100 p.p.t.v. and 10 p.p.b.v. (1 p.p.b.v. = 1 part in 10^9 by volume) was shown by the instrument. These latter standards were injected in 10-ml syringes so that the mass of F21 entering the detector at the 100 p.p.t. level was actually lower than that injected in the concentrated air samples, which typically contained the equivalent of 2 litres of air.

All of the data collected in this programme of F21 measurements are shown in Table 1. The first 14 results are values of the F21 concentration in the ambient air at Harwell, five of which used one technique for quantitative estimation and nine used another. The average value of the 14 estimations is 1.6 p.p.t.v. with $\sigma = 0.7$ p.p.t.v. There is no significant difference in the two groups of data obtained by the two techniques (1.8 ± 0.5 p.p.t.v. and 1.5 ± 0.7 p.p.t.v.), and this gives us confidence that our average of 1.6 p.p.t.v. is correct. The possibility that some compound other than F21 is being measured is not very high, as the results obtained using masses 69 and 102 are almost identical to those obtained using mass 67. Chloroethane, which contains a peak at m/e 67 in its mass spectrum, was observed to co-elute with F21 in this study; however, chloroethane does not have m/e peaks at 69 and 102.

Estimates of the F21 concentration in air samples from other locations are also shown in Table 1. Samples 16–19 represent

particularly clean air samples, and there the average falls to 0.08 p.p.t. Whether this number is significantly different from the noise level (~ 0.05 p.p.t.v.) is not possible to tell with the present experimental system. The numbers can be taken to indicate that the F21 concentration in the background troposphere is well below 1 p.p.t.v.

Tests have been made for contamination and decay in the containers used to bring the air to Harwell. Samples 15–18 were contained in aluminium cans with metal-bellows valves. There was no evidence for any decay in a similar can filled with Harwell air on 21 March 1979 and analysed on 21 May 1979 and 13 June 1979. Sample 19 was contained in a stainless steel can, but there is no reason to suspect decay in this system. Previous tests by one of us (R.A.R.) also showed that no growth of F21 occurred over a two-week period.

Our value for F21 in rural air in Southern England (1.6 p.p.t.v.) is substantially lower than previously reported estimates at other locations^{4–7}. Rasmussen¹⁰ has since revised his estimate of F21 concentrations originally reported (~ 10 p.p.t.v.) and now considers a value of 1.5 p.p.t.v. to be more likely. The data of Singh *et al.*^{6,7} were all obtained by GC/ECD and the chances of overestimation are quite large. A major problem in the analysis of F21 is to devise a column that will give complete separation of F21 from other neighbouring compounds. Rasmussen in his early work observed that chloroethane ($\text{C}_2\text{H}_5\text{Cl}$, boiling point, 12.7°C) had a retention time similar to F21 (boiling point, 8.9°C) in the conditions of analysis used with an SP 2100 column and temperature programming. Singh *et al.*'s¹¹ elution of F21 (boiling point, 8.9°C) before CH_3Br (boiling point, 3.5°C) on a DC 200 column substrate is inconsistent with the elution sequence of materials with increasing boiling points on that type of column. Additionally, the separation of F21 from CH_3I (boiling point, 43°C) on 0.5% SP 1000 Carbowax B column has presented difficulties¹². This leaves only Crescentini and Bruner's⁴ high value to be accounted for. Assuming that they may have been sampling polluted air, we collected air samples in the area between Liverpool and

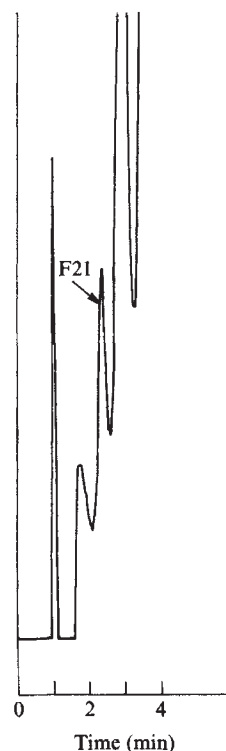


Fig. 1 F21 ambient air stored between 21 March and 21 May 1979; 2.6 p.p.t.v. at m/e 67. The sample size was 730 ml.

Table 1 Results of F21 measurement programme at Harwell

No.	Date	Sample	Vol. (ml)	F21 (p.p.t.v.)	Mass no.	Comments
1	7 November 1978	OA.1		1.9	67	COS used as internal standard
2		OA.2		2.6	67	
3		OA.3		2.2	67	
4	8 November 1978	OA.4		1.7	67	
5		OA.5		1.0	67	
6		OA.6	3 ml LA	1.1	67	
7	13 March 1979	OA.7	3 ml LA	1.1	69	Known volume of liquid air
8	14 March 1979	OA.8	3 ml LA	1.9	102	
9		OA.9	3 ml LA	2.3	102	
10		OA.10	3 ml LA	1.9	67	
11	21 May 1979	OA.11		1.1	67	
12		OA.12		0.5	67	
13		OA.13		0.6	67	
14	13 June 1979	OA.14		2.8	67	Known volume of liquid air
15	14 March 1979	Longview Al Plant (PNW)	3 ml LA	1.9	67	
16		Tasmania	3 ml LA	—	102	
17		Tasmania	3 ml LA	0.2	102	
18	15 March 1979	Background PNW	3 ml LA	0.07	67	
19		South Pole	3 ml LA	0.1	67	
20	21 May 1979	OA stored	3 ml LA	2.6	67	100 ml of gaseous air injected into cryotrap
21		21 March 1979	1 ml LA	2.6	67	
22			1 ml LA	2.0	67	
23	12 June 1979	Northwich	100 ml A	4.2	67	
24		Northwich	100 ml A	13	67	
25		Runcorn	100 ml A	37	67	
26		Runcorn	100 ml A	34	67	100 ml of gaseous air injected into cryotrap
27		OA	100 ml A	—	67	

OA, Outside air; LA, liquid air; A, air; PNW, Pacific North West; Al, aluminium.

Manchester. When measured using mass 67, samples 23–26 showed elevated levels of F21 with values between 4 and 37 p.p.t.v. Unfortunately these values were not confirmed by measuring masses 69 and 102, and samples collected more recently in this area show only very low F21 levels (~1.5 p.p.t.v.). Whilst it is thus quite feasible that Crescentini and Bruner were sampling F21 in polluted air, there still remains some doubt that concentrations as high as 20 p.p.t.v. occur.

If the only sink for F11 in the troposphere involved F21 production, then a lifetime for F11 could be derived from a knowledge of the relative F11/F21 concentrations and the rate of decay of F21. In the troposphere, the most likely sink for F21 is reaction with OH radicals, so:



and

$$\tau(\text{F11}) \sim \frac{[\text{F11}]}{[\text{F21}]} \tau(\text{F21})$$

This simplified equation yields a $\tau(\text{F11})$ of 200–300 yr. The concentrations of F11 and F21 were taken to be 165 p.p.t.v. and 1.6 p.p.t.v., respectively. $\tau(\text{F21}) \sim \{k[\text{OH}]\}^{-1}$ is between 2 and 3 yr based on the reaction rate constant of F21 with OH of $1.5 \times 10^{-12} \exp(-1,184/T) \text{ cm}^3 \text{ molecule}^{-1} \text{ s}^{-1}$ (ref. 13) at the average temperature of 260 K with mean OH densities in the range $0.6\text{--}1 \times 10^6 \text{ molecules cm}^{-3}$ (ref. 14). This equation for $\tau(\text{F11})$ is approximately valid when F11 and F21 are in a steady state. We calculated the lifetime of F11 using more complex expressions without the steady-state assumption, but the results were not significantly different. Furthermore, we used the average of the highest observed concentrations of F21 (1.6 p.p.t.v.). Our other measurements imply that, in fact, the global concentration of F21 is probably well below this value (~0.1 p.p.t.v.), which would in turn imply that such a conversion of F11→F21 is much less significant than calculated above. Therefore, our measurements show that the lifetime of F11 due to its possible conversion to F21 is much longer than that derived from stratospheric photolysis¹⁵, which is ~50 yr. For this tropospheric removal process of F11 to be equivalent to stratospheric photolysis, a background F21 concentration of 6 p.p.t.v.

would be required; a dominant tropospheric removal process would need a level exceeding 10 p.p.t.v.

In the discussion above, we have considered the extreme case where all the observed atmospheric burden of F21 results from destruction of F11. In fact, all or part of the atmospheric burden may be due to other sources which include direct industrial emissions⁸ and volcanic emissions^{16,17}.

Definite statements can be made only if more data are collected on the atmospheric concentrations of F21 in the background atmosphere, which would also justify more detailed theoretical treatments than discussed in this study. Our present experimental results show that non-urban F21 concentrations range from ~0.1 to 2.5 p.p.t.v., indicating that even if there are environmental processes which destroy F11, converting it to F21, they would not be a significant global, tropospheric sink for F11.

We thank Dr R. G. Derwent for helpful discussions and Steve Crawford and Bob Dalluge for laboratory assistance. This work was funded by the Chemical Manufacturers Association and the UK Department of the Environment. The samples 23–26 were obtained with help from ICI Ltd. Support from the NSF, grants ATM 7806628 and DPP 77-23468, is gratefully acknowledged.

Received 27 May; accepted 3 July 1980.

1. *Chlorofluorocarbons and their Effects on Stratospheric Ozone* (Pollut. Pap. No. 15, HMSO, London, 1979).
2. *Stratospheric Ozone, Depletion by Halocarbons, Chemistry and Transport* (National Academy of Sciences, Washington, DC, 1979).
3. Ausloos, P., Rebert, E. R. & Glasgow, L. C. *Natn. Bur. Stand. J. Res.* **82**, 1 (1977).
4. Crescentini, G. & Bruner, F. *Nature* **279**, 311 (1979).
5. Rasmussen, R. A. *EPA Rep. #R0804033-01* (1976).
6. Singh, H. B., Salas, L., Shigeishi, H. & Crawford, A. *Atmos. Envir.* **11**, 819 (1977).
7. Singh, H. B., Salas, L., Shigeishi, H. & Scribner, E. *Science* **203**, 899 (1979).
8. Glasgow, L. C., Jesson, J. P. & Ward, R. B. *Atmos. Envir.* **12**, 692 (1978).
9. Rasmussen, R. A., Penkett, S. A. & Prosser, N. *Nature* **277**, 549 (1979).
10. Rasmussen, R. A. *Atmospheric Measurement of CHFCl₂ (F21)*, Final Rep. No. 78-60, Chemical Manufacturers Association, London, 1979).
11. Singh, H. B., Salas, L. J., Shigeishi, H. & Smith, A. H. *Fate of Halogenated Compounds in the Atmosphere Interim Rep.* No. PB 278-198 to EPA, 19 (1978).
12. Bruner, F., Bertoni, G. & Crescentini, G. *J. Chromatogr.* **167**, 399 (1978).
13. DeMore, W. B. *Chemical Kinetics and Photochemical Data for Use in Stratospheric Modelling Evaluation Numbers* (Jet Propulsion Laboratory, Pasadena, 1979).
14. Volz, A., Ehhalt, D. H., Derwent, R. G. & Khedim, A. *Ber. der Kernforsch.* No. 1604 (1979).
15. Derwent, R. G. & Eggleton, A. E. *J. Harwell Rep. AERER 9112* (HMSO, London, 1978).
16. Stoiber, R. E., Leggett, D. C., Jenkins, T. F., Murrmann, R. P. & Rose, W. I. *Bull. geol. Soc. Am.* **82**, 2299 (1971).
17. Graedel, T. E. *Chemical Compounds in the Atmosphere* (Academic, New York, 1978).

A climatically responsive tree-ring record from *Widdringtonia cedarbergensis*, Cape Province, South Africa

Peter W. Dunwiddie

Quaternary Research Center, University of Washington, Seattle, Washington 98195

Valmore C. LaMarche Jr

Laboratory of Tree-Ring Research, University of Arizona, Tucson, Arizona 85721

The use of dated tree-ring series to reconstruct past climate has received much attention in recent years^{1,2}. The development of long and climatically responsive chronologies in many areas of the Southern Hemisphere is an important step towards better understanding past environments in these areas^{3,4}, but little dendrochronologically related work of any sort has been done on the African continent⁵⁻¹¹, and most of this has consisted primarily of ring counts to estimate ages. We report here on the development and climatic interpretation of the first dated annual ring-width index chronology from Africa south of the Sahara.

Tree-ring samples were collected from a large number of sites in South Africa¹² as part of a broader-scale effort to reconstruct climatic variability in the Southern Hemisphere. The results of reconnaissance sampling, combined with previous experience elsewhere, led us to concentrate on conifers in the genera *Podocarpus* (Podocarpaceae) and *Widdringtonia* (Cupressaceae) in western and southern Cape Province. One species, *W. cedarbergensis*, seemed to offer particular promise by virtue of its apparent great age, its tree-ring properties, and its occurrence in marginal habitats of the Cedarberg Mountains.

The Die Bos site, which provided the tree-ring samples on which this analysis is based, is located in the rocky Cedarberg Mountains about 200 km north of Cape Town. It is one of the

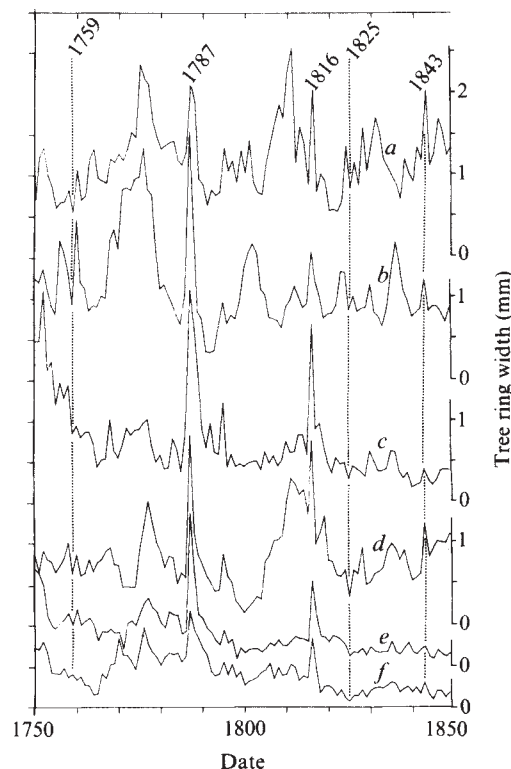


Fig. 2 Portions of ring-width plots for six trees from the Die Bos site illustrating crossdating in the period 1750-1850.

largest remaining stands of *W. cedarbergensis*, whose numbers have been greatly reduced by woodcutting and wild fires. Scattered *Protea* (sp.) trees and shrubs are the main associates on this sparsely vegetated site. The climate in the Cedarberg is moderately warm in summer and cold in winter, and the area receives between 300 and 900 mm of precipitation per year, mainly in winter.

Our sample includes 58 increment cores from 28 living trees up to 12 m tall and 150 cm in diameter. These are supplemented by cross-sectional disks from 19 dead trees and stumps. Lobate growth is uncommon in this material, as is the wedging out of individual rings. Circuit uniformity is generally good, with ring-width patterns being similar on various radii of the same tree. However, several growth features make cross-dating unlike that in most other species (Fig. 1). Many rings lack an abrupt termination of latewood growth, and may have several false bands within the latewood, making determination of the precise ring boundary difficult. Resin-filled bands of cells occur in some rings that may appear as sharper, more distinct lines than the true boundary. The 'signature' rings commonly used by dendrochronologists to crossdate specimens (rings that are relatively narrow on many trees in a site) are not common. Rather, unusually wide rings (Fig. 2), which in some cases (1787, 1816) exhibit characteristic frost damage in the earlywood are more reliable dating checks. These frost rings (Fig. 1) are unusual because trees of most species are not susceptible to such damage after they reach several centimetres in diameter. At Die Bos, mature trees are found with frost rings that date consistently between trees. It is not known why these rings are often relatively wide as well. Figure 2 also illustrates the wide variation in growth rates and sensitivities found in trees at this site. Some, like trees *c* and *e*, are relatively slow-growing, but have several extremely wide rings. Others, such as *a* and *b*, have many relatively wide rings, some of which crossdate with other trees. There is little doubt that the rings dated in this chronology are annual, for several reasons. First, the extreme seasonality of the

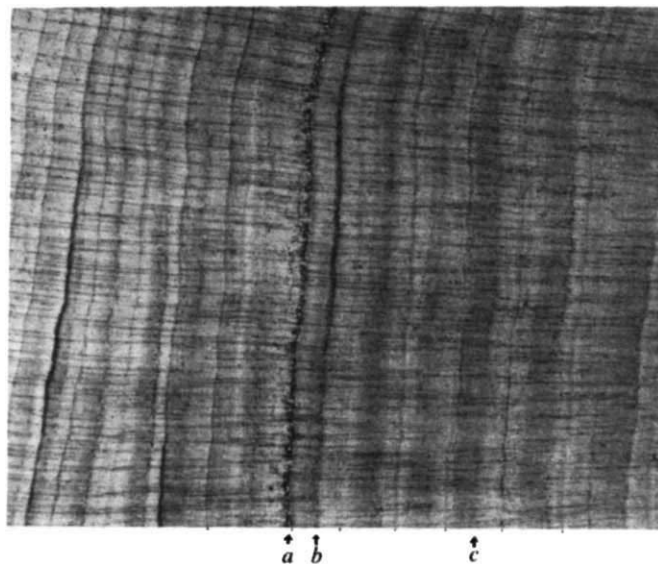


Fig. 1 Detail of growth rings in *W. cedarbergensis*. *a*, Frost ring in earlywood of 1787; *b*, resin band in centre of 1786 ring; *c*, intra-annual band in latewood of 1783 ring. Rings following 1787 have clearer latewood terminations than do earlier ones.

Mediterranean-type climate is likely to limit growth to a single growing season each year. Second, locally absent ('missing') rings are infrequent, and relatively few rings in any specimen are ambiguous; false bands can usually be clearly differentiated from the annual rings by careful inspection and crossdating.

The site chronology beginning in AD 1564 (Fig. 3) was developed by standard methods¹ after all suitable samples were dated. It represents the average of ring-width index series from 52 radii on a total of 32 trees, an unusually large sample size. An analysis of variance¹ based on data for the 1892–1976 subperiod showed that 29% of the total sample variance is retained by the site chronology, indicating a fairly high degree of ring-width variation in common between different radii and different trees. Presumably, this reflects a fair degree of common response to year-to-year variations in the local environment, including climate.

A climatic response function¹ was computed to evaluate the nature of the climatic 'signal' contained in the site chronology (Fig. 4). It is based on comparison, using multivariate analytical methods, of monthly climatological data from a nearby station with the annual ring-width index during the period 1931–60. The climatic variables used are the monthly average daily maximum temperature and the total monthly precipitation during each month in an 18-month climatic interval beginning in spring (October) of the year preceding initiation of ring growth, and ending in autumn (March) of the calendar year following the year in which growth begins. That is, this interval includes the past and present growing seasons. The growing season for this species probably includes the period October through March. Previous ring-width index was not used in generating the response function shown here, but incorporating previous growth in the analysis produces very similar results.

The response function results show a high degree of correlation of tree growth with climatic factors; 73% of the year-to-year variability in ring-width index can be explained by departures from normal temperature and precipitation during the 18-month interval preceding and concurrent with the growing season. However, the response is complex, with opposite responses to the same variables occurring in different seasons of the year. A dominant feature is the positive response to precipitation of the previous and current spring and early summer (September or October to December). Coupled with a negative response to higher than normal temperatures in the previous and current early summer (November–December), this suggests that one component of the relationship is a negative response of tree growth to spring and early summer drought. However, in some months, a negative response to precipitation and positive response to temperature also occur. We can tentatively interpret the long tree-ring chronology (Fig. 3) as a record of spring and early summer moisture availability, but other chronologies from a greater diversity of habitats and a broader geographic area must be developed, studied, and compared to understand better the climatic history of southwestern South Africa over the past several hundred years. For example, the occurrence of frost rings in most trees following the extremely large volcanic eruption of Tambora at 8°S in 1815 may suggest a dust veil effect on temperatures of interest to climatologists.

Dendrochronological study of *W. cedarbergensis* can also provide a fire frequency record. Several charred disks from the

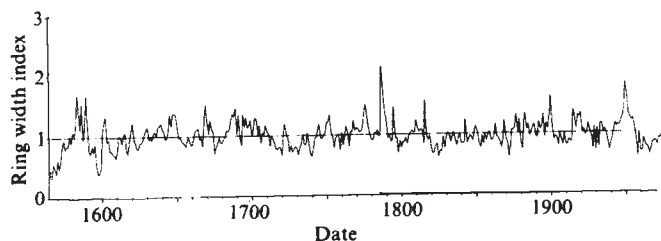


Fig. 3 Site chronology of average ring-width indices, Die Bos, Cape Province.

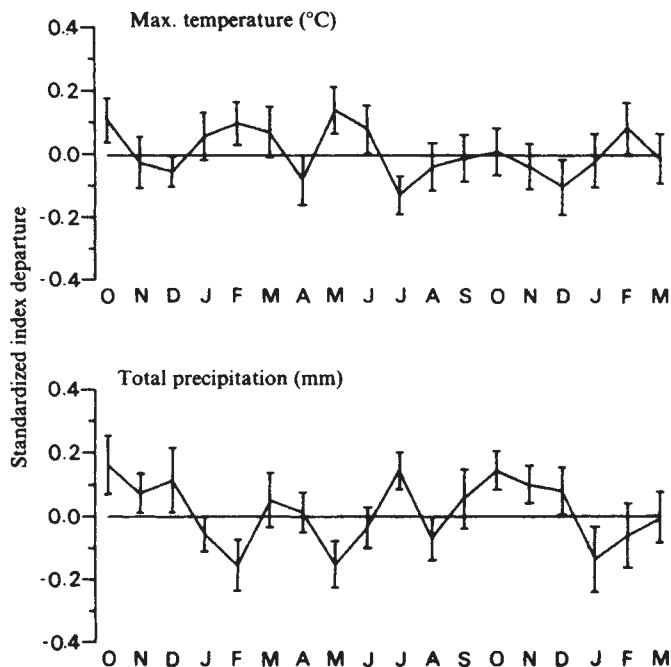


Fig. 4 Climatic response function for the Die Bos *Widdringtonia cedarbergensis* chronology. Based on combined climatological data from Clanwilliam, Calvinia, and Langevans, Cape Province, 1931–60.

Die Bos site have an outer ring dated 1872, and some living trees exhibit a marked growth surge beginning the next year, suggesting a release effect following fire. A marked resin band in the ring for 1771, along which many specimens split, may be the result of an earlier fire. Fire-scarred living trees are unusual in this resinous, flammable species; trees are generally either killed or left unscarred by fire.

There is potential for development of much longer tree-ring chronologies than that presented here. Dead wood decays slowly in this semi-arid environment, especially in rock crevices or overhanging rock shelters, and wood from stumps, logs, and remnants can be cross-dated with living material. Ring-width plots for trees *a* and *d* (Fig. 2) illustrate the matching of such 'floating' series with growth records of living trees at the site.

This analysis illustrates the potential for development and climatic interpretation of long tree-ring records from southern Africa. In addition to study of ring-width variability, other approaches such as analysis of stable isotopic composition may yield further information. Some of this material has already been supplied for isotopic analysis.

We thank F. Kruger, A. Lamb, and F. Hanekom for support and field assistance, E. Ashton for selected climatological data from copies of the South African Weather Bureau's magnetic tapes. This research was supported by the Climate Dynamics Program, Division of Atmospheric Sciences, NSF.

Received 11 April; accepted 5 May 1980.

1. Fritts, H. C. *Tree-Rings and Climate* (Academic, London, 1976).
2. LaMarche, V. C. Jr *Nature* **276**, 334–38 (1978).
3. Dunwiddie, P. W. *N.Z.J. Bot.* **17**, 251–66 (1979).
4. Ogden, J. *Aust. J. Ecol.* **3**, 339–56 (1978).
5. Walter, V. H. *Naturwissenschaften* **28**, 607–12 (1940).
6. Guy, G. L. *Proc. Trans. Rhod. Sci. Ass.* **54**, 68–84 (1970).
7. Storry, J. G. S. *Afr. J. Sci.* **71**, 248 (1975).
8. Hall, M. *Ann. Natal Mus.* **22**, 693–703 (1976).
9. Curtis, B. A., Tyson, P. D. & Dyer, T. G. J. S. *Afr. J. Sci.* **74**, 92–95 (1978).
10. Lilly, M. A. *An assessment of the dendrochronological potential of indigenous species of trees in South Africa*, Occ. Paper No. 18 (Dept Geogr. & envir. Stud., Univ. of Witwatersrand, Johannesburg, 1977).
11. Munaut, A. V., Berger, A. J., Guiot, J. & Mathieu, L. *Evolution des Atmosphères Planétaires et Climatologie de la Terre* (Centre National d'Etudes Spatiales, Nice, 1978).
12. LaMarche, V. C. Jr, Holmes, R. L., Dunwiddie, P. W. & Drew, L. G. *Laboratory of Tree-Ring Research Chronology Series 5*, Vol. 5 (University of Arizona, Tucson, 1979).

Organic matter fluxes from sediment traps in the equatorial Atlantic Ocean

Stuart G. Wakeham, John W. Farrington, Robert B. Gagosian, Cindy Lee, Hein DeBaar, Gale E. Nigrelli, Bruce W. Tripp, Steven O. Smith & Nelson M. Frew

Department of Chemistry, Woods Hole Oceanographic Institution, Woods Hole, Massachusetts 02543

The vertical flux of particulate organic matter to the ocean floor is controlled by complex remineralization and transport processes. Rapidly sinking, large ($>32\ \mu\text{m}$) particles may account for most of the vertical mass flux¹⁻³. Experiments involving collection of sedimenting particles in traps deployed at varying depths in the water column provide a way of assessing these processes directly²⁻¹². Measurement of organic carbon content of trap material may lead to a general understanding of the flux of particulate organic matter to the sea floor^{3,8-10}. However, details of the transport and transformation phenomena involving sinking particulate organic material can be elucidated only by determining distributions of specific organic compounds associated with the particles^{11,12}. We report here direct flux measurements and composition of lipids obtained by organic geochemical studies of particulate material collected in a deep-sea moored sediment trap experiment in the equatorial Atlantic Ocean.

The PARFLUX E station was located on the Demerara abyssal plain (5,288 m) ~750 km from the Guyana coast at 13°30.2' N, 54°00.1' W. Four sediment traps were placed along a fixed mooring at depths of 389, 988, 3,755 and 5,068 m, and recovered after 98 days (November 1977–February 1978). The water column had a weak oxygen minimum between 200 and 400 m, and the main thermocline was located at 300 m. No significant nepheloid layer was reported⁷. Sodium azide was added to each trap before emplacement and diffused into the collection cup¹³ during the experiment to retard bacterial spoilage of the trapped material.

Particulate samples were wet sieved and split with a rotary splitter³. Aliquots were taken for geological³, elemental and isotopic⁷, and organic geochemical (present) studies. About one-sixteenth of each sample (0.03–0.4 g sediment depending on availability) was extracted with toluene/methanol (1:1). Amino acids present in this extract were partitioned into saturated salt solution, while the remaining lipids were partitioned into hexane. The hydrolysed amino acids were analysed by a modification of the fluorescence technique of Hill *et al.*¹⁴. An aliquot of the hexane-soluble lipids (0.5–5 mg of lipid) was hydrolysed for total fatty acid analysis. A second aliquot (0.4–4 mg) was separated into constituent organic compound classes by adsorption chromatography (Merck silica gel 60, 5% deactivated with water) by elution with a series of solvents of increasing polarity (ranging from hexane, toluene in hexane, increasing proportions of ethyl acetate in hexane, ethyl acetate, to methanol).

The molecular composition of compound classes collected (given in Table 1) was determined by glass capillary gas chromatography (GC) (Carlo Erba 2150 and 4160, and Hewlett Packard 5710 and 5840 chromatographs; 15–25 m $\times$ 0.3 mm i.d. columns, deactivated by persilylation and coated with SE-52¹⁵) and capillary gas chromatography/mass spectrometry (GC/MS) (Finnigan 3200 and 1015 SL). Electron impact and isobutane- or methane-chemical ionization spectra were obtained, depending on the compound class. Fatty acids were analysed as methyl esters¹⁶, and fatty alcohols and sterols as acetate derivatives¹⁷. Wax esters, steryl esters, and triacylglycerols were analysed intact by capillary GC and GC/MS.

Compounds were identified where possible by comparison with GC retention times and mass spectra of authentic standards. Concentrations of hexane-soluble lipids were estimated by weighing an aliquot on an electrobalance; concentrations for compound classes are sums of individual compounds determined by gas chromatography. Laboratory blanks were carried through the entire workup and were generally significantly lower than samples. We cannot estimate sampling variability as single samples were collected, but analytical precision for each compound class is better than $\pm 20\%$.

Organic geochemical results for PARFLUX E sediment trap samples are given in Table 1 as total amounts of material in each trap and as mass fluxes at the depths sampled. The major compositional features of the lipid classes analysed are summarized in Table 2. Further details of compound distributions will be given elsewhere.

Particles $<1\ \text{mm}$ in size accounted for most of the particulate flux at all depths sampled. The amount of material $<1\ \text{mm}$ varies little with depth in the water column. Honjo³ found these particles to be amorphous aggregates of organic matter composed of foraminiferal tests, radiolarian skeletons, small pteropod shells, and faecal pellets. Only in the 389 m trap were particles $>1\ \text{mm}$ important, these including larger faecal pellets and some intact zooplankton³.

Bulk organic matter flux decreased rapidly with increasing depth in the mesopelagic zone (389 and 988 m traps), stabilizing in the bathypelagic layer (3,755 and 5,068 m samples). Organic carbon comprised ~50% of the organic matter in the mesopelagic samples and ~35% at the two deeper depths. Of the organic carbon fixed in the euphotic zone at this location³ ($200\ \text{mg m}^{-2}\ \text{day}^{-1}$), ~4–6% reached the mesopelagic traps. Only ~1% arrived in the deep traps. These observations are consistent with a rapid turnover of bulk organic matter in the upper portion of the water column, with a small and highly fractionated amount reaching the sea floor.

Lipids were rapidly depleted in the upper 1,000 m of the water column. Unlike bulk organic matter and organic carbon, lipid fluxes decreased throughout the water column. Lipids became a proportionally smaller fraction of the organic flux (15% at 389 m and 2% at 5,068 m), indicating that they were consumed preferentially compared with other components of the overall organic pool.

Fluxes for specific organic compound classes followed three trends (Table 1): (1) a rapid decrease with depth (hydrocarbons, fatty acids, wax esters, steroid ketones, and lipid-associated amino acids); (2) a rapid decrease between 389 and 988 m followed by a slight flux increase in the 3,755 m trap (free fatty alcohols); (3) a maximum flux at 988 m (free sterols, steryl esters, triacylglycerols). Major aspects of our observations are outlined below and in the tables.

Hydrocarbons of planktonic origin ($n\text{-C}_{17}$, C_{19} , and a lesser amount of pristane) dominated all four traps. Long-chain (C_{25} – C_{31}) hydrocarbons were also present at lower abundances at all depths, indicating long-range transport of continentally derived material. Sterenes were detected, primarily in the 389 and 988 m samples, where they were more abundant than alkanes in the same carbon range. This suggests active microbial transformations of organic matter in or on sinking particles and a relatively fast biogeochemical formation for sterenes¹⁸. We cannot, however, rule out a planktonic source for these cyclic hydrocarbons.

The flux of total fatty acids (free + esterified) decreased 60-fold with increasing depth. Free fatty acids, a minor fraction of the total acid pool, showed a smaller decrease. The fatty acid composition of the particulates reflects a planktonic source. Recycling and sinking of particles resulted in rapid degradation of all fatty acid moieties, unsaturated compounds being lost more rapidly than saturated ones (for example, saturated/unsaturated acids = 5 and 60 at 389 and 5,068 m respectively). Branched-chain fatty acids, potential markers of microbial contribution¹⁹⁻²², were minor constituents, and their relative abundance did not increase with depth.

Table 1 Organic geochemical results for PARFLUX E sediment traps

Depth of 1.5 m ² trap (m)	Sample designation	Particle size	Particulate matter*	Organic matter†	Organic carbon†	Hexane-soluble lipids‡	Hydrocarbons§	Total fatty acids§	Free fatty acids§	Wax esters§	Steroid ketones§	Amino acids (Lipid fraction)§	Free fatty alcohols§	Free sterols§	Steryl esters§	Triacylglycerols§
			(g)			(mg)	Amount of material per trap									
							(µg)									
389	E6	<1 mm	6.6	1.4	0.7	225	1,200	39,100	4,100	770	4,720	300	20,500	19,100	170	390
	E2	>1 mm	3.6	0.6	0.3	70	250	16,500	2,600	530	130	90	4,100	7,400	60	90
988	E3	<1 mm	6.7	1.2	0.5	170	1,120	17,400	470	180	3,300	43	1,800	31,400	2,300	2,100
	E1	>1 mm	0.5	0.1	0.04	20	170	5,100	—	20	20	3.2	980	3,360	200	110
3,755	E4	<1 mm	6.8	0.7	0.2	25	140	1,280	—	16	530	13	6,300	900	8	17
5,068	E5	<1 mm	6.9	0.7	0.2	13	97	910	200	49	590	14	760	550	9	12
							Fluxes									
			(mg m ⁻² day ⁻¹)				(µg m ⁻² day ⁻¹)									
389	E6	<1 mm	45.1	9.7	4.7	1,530	8.2	270	27.9	5.3	32.1	2.0	140	130	1.2	2.6
	E2	>1 mm	24.2	3.9	2.0	470	1.7	110	17.8	3.6	0.9	0.6	30	50	0.4	0.6
988	E3	<1 mm	45.9	7.9	3.7	1,160	7.6	120	3.2	1.2	22.4	0.3	12	214	15.9	14.3
	E1	>1 mm	3.3	0.7	0.3	130	1.2	35	—	0.1	0.1	0.02	7	23	1.3	0.7
3,755	E4	<1 mm	46.4	4.8	1.7	170	1.0	8.7	—	0.1	3.6	0.1	43	6	0.06	0.12
5,068	E5	<1 mm	47.0	4.9	1.7	85	0.7	6.2	1.4	0.3	4.0	0.1	5	4	0.06	0.08

* Weight calculated from weighing replicate aliquots of splits³.

† Data from ref. 3.

‡ Gravimetric determination.

§ Determined by GC or HPLC (amino acids).

|| Not detectable relative to blank.

Unsaturated wax esters (for example, $C_{34:1}/C_{34:0}=40$ at 389 m) dominated the wax ester flux. The composition was indicative of a zooplankton origin^{23,24}. Degradation of unsaturated wax esters appeared faster than loss of saturated analogues. However, even at 5,068 m depth, unsaturated compounds dominated the wax ester distribution ($C_{34:1}/C_{34:0}=15$).

Steroid ketones (stanones + stenones) exhibited a gradual flux decrease with increasing depth. The major ketone in each sample was cholest-4-en-3-one, representing 60% of the ketone flux into the mesopelagic traps and about 35% into the deeper traps. *In situ* production of steroid ketones, possibly through microbial transformations of sterols, may have occurred as the particles sank. The gradual flux decrease of ketones compared to sterols (see below) could be due to differing degradation rates.

Amino acids associated with the lipid fraction in the mesopelagic traps were similar in composition to the protein amino acid composition of plankton. Such distributions were also found in suspended particulate matter filtered from seawater in the Pacific²⁵. In contrast, only glutamic acid, aspartic acid, alanine, serine and glycine were detected in the deeper traps. These are the major amino acids in bacterial cell walls²⁶ and may reflect bacterial input.

Free fatty alcohols of 14, 16 and 18 carbon atoms were associated with particles in all four traps. Only about 10% of the alcohols in the 389 m trap were found in the 988 m trap. On the other hand, we observed a significant enrichment of fatty alcohols at 3,755 m. The several compounds responsible for this increase remain unidentified at present, but they are neither straight-chained molecules nor phytol.

The free sterol flux maximized at 988 m, where cholest-5-en-3 β -ol accounted for 95% of the sterol composition. In contrast, at 389, 3,755 and 5,068 m, a wide range of planktonic sterols were present.

Steryl ester fluxes and molecular distributions correlate well with free sterols, although the relative flux increase at 988 m over 389 m was about an order of magnitude greater for the esters. In the 389 m sample, steryl esters were composed of a

range of sterols (Table 2) and primarily $C_{16:0}$ and $C_{18:0}$ fatty acids. However, esters of cholest-5-en-3 β -ol predominated at 988 m.

Triacylglycerols in the 389 m trap were composed primarily of $C_{18:0}$, $C_{18:1}$ and $C_{16:0}$ fatty acids, resulting in a C_{45} - C_{63} carbon range. Polyunsaturated fatty acid moieties ($C_{22:6}$ and $C_{20:5}$) were more abundant in the 988 m sample than saturated or monounsaturated acids.

Our results suggest that significant alteration of organic compounds produced by phytoplankton and zooplankton^{21,23,24,27-29} in the upper water column takes place even before the particles reached the shallowest trap. For example, n - C_{17} , C_{19} , and pristane were depleted relative to other alkanes in the 389 m trap compared with freshly sampled indigenous zooplankton collected at 300 m at the PARFLUX E site (our analyses). Unsaturated fatty acids and wax esters in this trap were apparently degraded more rapidly than saturated acids and esters. On the other hand, sterenes and steroid ketones were generally absent from the plankton lipid but were abundant in the trap samples.

Details of food web dynamics, microbial transformations and biochemical reactions in the open ocean are poorly understood. However, it is apparent that many of the organic compounds we analysed are consumed (recycled and/or remineralized) in the mesopelagic layer in this part of the ocean. Only a few per cent of these compounds are associated with particles reaching the bathypelagic zone. Nevertheless, alterations of the composition of extractable organic matter on the sinking particles occur throughout the water column and significant amounts of labile organic material reach the sea floor. Many of the compositional changes we observed take place below the ~300 m depth below which Menzel³⁰ thought most organic compounds were refractory.

We cannot preclude that vertically migrating zooplankton were not entrapped in the mesopelagic sediment traps. However, it is unlikely that the flux maxima at 988 m for sterols, steryl esters, and triacylglycerols are due to such entrapment. Only in the 389 m trap did particles >1 mm contribute

Table 2 Composition of lipid classes in PARFLUX E sediment traps

Compound class	Range	Compound distribution
Hydrocarbons	C ₁₅ –C ₃₂	<i>n</i> -C ₁₇ , C ₁₉ dominate all depths; CPI ₂₃₋₂₉ = 1.2–1.4 sterenes (such as, cholest-2-ene, 3,5-cholestadiene) abundant
Fatty acids	C ₁₄ –C ₂₆	C _{16:0} *, C _{18:1} , C _{18:0} dominate all depths† C _{20:5} , C _{22:6} —10–20% of total acids in mesopelagic traps, 1–2% in bathypelagic traps Iso, anteiso C ₁₅ , C ₁₇ minor (<2%) at all depths
Wax esters	C ₂₆ –C ₄₄	C _{34:1} , C _{32:1} dominate all depths (~75% of total wax ester at 389 m; ~50% at 988 m and 5,026 m; 25% at 3,755 m depth) Fatty acids: C ₁₄ –C ₂₆ ; C _{16:0} , C _{22:0} , C _{18:1} major Fatty alcohols: C ₁₄ –C ₂₆ ; C _{16:0} , C _{18:1} , C _{16:1} , C _{18:0} major
Steroid ketones	—	Cholest-4-en-3-one dominates (>60% of total ketones in mesopelagic traps, ~35% in bathypelagic traps)
Amino acids (extractable)	—	Typical planktonic amino acids in upper traps Glutamic acid, aspartic acid, alanine, serine, glycine only in deeper traps suggesting bacterial input (?)
Free fatty alcohols	C ₁₂ –C ₂₆	C _{16:0} , C _{18:1} , C _{18:0} , C _{16:1} dominate
Free sterols	—	Cholest-5-en-3 β -ol dominant sterol at all depths (95% of sterol distribution at 988 m; 20–50% in other depths) Cholesta-5,22(E)-dien-3 β -ol, 24-methylcholesta-5,22(E)-dien-3 β -ol, 4,23,24-trimethyl-5 α -cholest-22-en-3 β -ol abundant
Steryl esters	—	Fatty acids: 389 m C _{16:0} , C _{18:0} major 988 m C _{16:0} , C _{14:0} , C _{18:0} , C _{18:1} major Sterols: 389 m cholest-5-en-3 β -ol (60%), cholesta-5,22(E)-dien-3 β -ol, 24-methylcholesta-5,22(E)-dien-3 β -ol, 5 α -cholestan-3 β -ol 988 m cholest-5-en-3 β -ol (85%), 5 α -cholestan-3 β -ol, cholesta-5,22(E)-dien-3 β -ol
Triacylglycerols	C ₄₅ –C ₆₃ ‡	C ₅₁ , C ₅₃ , C ₅₅ , C ₅₇ compounds predominate Fatty acids: 389 m C _{18:0} , C _{18:1} , C _{16:0} major 988 m C _{22:6} , C _{20:5} also dominant

* Shorthand notation denotes number of carbon atoms: number of double bonds.

† For each category, compounds are listed in order of decreasing relative abundances.

‡ Total carbon number (acyl carbons + glycerol carbons).

significantly to the flux of organic compounds (Table 1). Furthermore, at 988 m the individual molecular constituents of the sterol, steryl ester, and triacylglycerol fractions were vastly different compared with both shallower and deeper samples (Table 2). Two hypotheses may be proposed to account for these maxima. First, differences in biological communities populating this region of the water column may be responsible for the different inputs of these organic compound classes at depth compared to shallower than 389 m. Alternatively, organic matter produced in the euphotic zone may be compartmentalized and transported through 389 m, followed by the release of these compounds below the 389 m trap.

Our interpretations are constrained by the limited data set—a single set of samples from one location in the World's ocean. In addition, our lipid analyses account for only ~15–20% of the total organic flux. However, we are pursuing a better understanding of the cycling of organic matter in the oceans by continuing detailed organic analyses of particulate matter collected in sediment traps in both open ocean (North central Pacific) and coastal (Peru upwelling, Panama Basin, and coastal California) areas. Interpretation of all sediment trap results are complicated by questions of differential settling rates, resuspension of bottom sediments and horizontal transport processes,

trapping efficiencies, and the potential for chemical and biochemical reactions after particles have been intercepted.

Several authors have focused on the measurement of total organic carbon associated with particles caught in sediment traps^{2-5,8-10}. Our results, however, show for the first time that while total organic carbon fluxes in the deep ocean decrease only slightly throughout the water column, the composition of portions of this material changes dramatically. These changes reflect processes acting on the particles during transport through the water column, and they are important to the composition of the organic matter being deposited at the sediment–water interface.

This research was supported by contract NOOO14-79C-0071 from the Office of Naval Research and grants OCE 78-26084, OCE 78-26180, and OCE 77-27004 from the National Science Foundation. We thank Susumu Honjo, Peter G. Brewer and Derek W. Spencer for sediment trap samples, and Carl O. Erba for analytical assistance and S. Henrichs, D. Repeta, J. Hunt and S. Honjo for valuable comments. Woods Hole Oceanographic Institution contribution No. 4569.

Received 19 March; accepted 9 June 1980.

- McCave, I. N. *Deep-Sea Res.* **22**, 491–502 (1975).
- Honjo, S. *J. mar. Res.* **36**, 469–492 (1978).
- Honjo, S. *J. mar. Res.* **38**, 53–97 (1980).
- Berger, W. H. & Soutar, A. *Science* **156**, 1495–1497 (1967).
- Wiebe, P. H., Boyd, S. H. & Winget, C. *J. mar. Res.* **34**, 341–354 (1976).
- Spencer, D. W. *et al. J. mar. Res.* **36**, 493–523 (1978).
- Brewer, P. G., Nozaki, Y., Spencer, D. W. & Fleer, A. P. *J. mar. Res.* (in the press).
- Knauer, G. A., Martin, J. H. & Bruland, K. W. *Deep-Sea Res.* **26**, 97–108 (1979).
- Hinga, K. R., Sieburth, J. McN. & Heath, J. R. *J. mar. Res.* **37**, 557–579 (1979).
- Deuser, W. G. & Ross, E. H. *Nature* **283**, 364–365 (1980).
- Crisp, P. T., Brenner, S., Venkatesan, M. I., Ruth, E. & Kaplan, I. R. *Geochim. cosmochim. Acta* **43**, 1791–1801 (1979).
- Prahl, F. G. & Carpenter, R. *Geochim. cosmochim. Acta* **43**, 1959–1972 (1979).
- Honjo, S. & Connell, J. F. *Tech. Rept.*, 79–80 (Woods Hole Oceanographic Institution, 1979).
- Hill, D. W., Walters, F. H., Wilson, T. D. & Stuart, J. D. *Analyt. Chem.* **51**, 1338–1341 (1979).
- Grob, K., Grob, G. & Grob, K. Jr *H.R.C. & C. C.* **2**, 31–35 (1979); **2**, 677–678 (1979).
- Metcalf, L. D. & Schmitz, A. A. *Analyt. Chem.* **33**, 363–364 (1963).
- Galli, G. & Paoletti, E. G. *Lipids* **2**, 72–75 (1967).
- Gagosian, R. B. & Farrington, J. W. *Geochim. cosmochim. Acta* **42**, 1091–1101 (1978).
- Leo, R. F. & Parker, P. L. *Science* **152**, 649–650 (1966).
- Cooper, W. J. & Blumer, M. *Deep-Sea Res.* **15**, 535–540 (1968).
- Oliver, J. D. & Colwell, R. R. *Int. J. Syst. Bact.* **23**, 442–458 (1973).
- Perry, G. J., Volkman, J. K., Johns, R. B. & Bauer, J. H. Jr *Geochim. cosmochim. Acta* **43**, 1715–1725 (1979).
- Nevenzel, J. C. *Lipids* **5**, 308–319 (1970).
- Sargent, J. R., Lee, R. F. & Nevenzel, J. C. in *Chemistry and Biochemistry of Natural Waxes* (ed. Kolattukudy, P. E.) 49–91 (Elsevier, New York, 1976).
- Siezen, R. J. & Magee, T. H. *Mar. Chem.* **6**, 215–231 (1978).
- Salton, M. R. J. *Microbial Cell Walls* (Wiley, New York, 1961).
- Goad, L. J. in *Marine Natural Products: Chemical and Biological Perspectives* Vol. 2 (ed. Scheuer, P. J.) 76–172 (Academic, London, 1976).
- Sargent, J. R. in *Biochemical and Biophysical Perspectives in Marine Biology* (eds Malins, D. C. & Sargent, J. R.) 149–212 (Academic, London, 1976).
- Schmitz, F. J. in *Marine Natural Products: Chemical and Biological Perspectives* Vol. 1 (ed. Scheuer, P. J.) 241–297 (Academic, New York, 1973).
- Menzel, D. W. in *The Sea*, Vol. 5 (ed. Goldberg, E. D.) 659–678 (Wiley, New York, 1974).

First age-diagnostic fossils from the central part of the North Greenland foldbelt

Finn Surlyk*, John Malcolm Hurst† & Merete Bjerreskov‡

*Geologisk Museum, Øster Voldgade 5, 1350 København K, Denmark
†Grønlands Geologiske Undersøgelse, Øster Voldgade 10, 1350 København K, Denmark

‡Institut for Historisk Geologi og Palaeontologi, Øster Voldgade 10, 1350 København K, Denmark

The discovery of Lower Llandovery (Silurian) graptolites in Peary Land, near the top of the sequence in the central part of the North Greenland foldbelt, shows that the rocks in this sequence are mostly pre-Silurian. It also permits correlation of the lithostratigraphic units in the central foldbelt¹ with those on the margin² and on the unfolded platform to the south^{3,4}. This places constraints on the timing of the Innuitian orogeny in North Greenland.

A 600 km long east-west trending Palaeozoic foldbelt forms the Arctic coastal area of North Greenland³ (Fig. 1). The foldbelt continues to the west as the Innuian mobile belt of the Canadian Arctic islands⁵. In Greenland the belt is up to 100 km wide and the degree of deformation and metamorphism increases from south to north¹. The folded sequence includes Upper Precambrian to Lower Palaeozoic rocks of a northern clastic trough and the adjacent part of a southern carbonate platform.

The evolution of the platform-trough system started in the late Precambrian with the development of an incipient east-west trending trough. Faulting and subsidence continued through Cambrian, Ordovician and Silurian times. The trough was formed by rifting and the southern margin was controlled by major faults. The northern margin is not known in Greenland as it was probably situated under the present Arctic Ocean. The trough expanded in several episodes by southwards shift of the southern margin to new fault systems. During this process, the older fault systems became intrabasinal and largely inactive (Fig. 2).

The first Precambrian?–Cambrian phase of trough fill mainly comprised turbidites, dark mudstones, light lime mudstones and resedimented conglomerates (quartzite and Paradisfjeld Groups; Fig. 3) reflecting the incipient nature of the rift. Subsequently, Cambrian turbidite sedimentation reflects deeper water and a greater differentiation of shelf and trough. Transport was both lateral from south and north and longitudinal from the east. At the end of this phase turbidite sedimentation was punctuated by several episodes of conglomeratic gravity flows initiated on the platform margin and upper slope. The turbidites and conglomerates represent the Polkorridoren Group¹ and formation A (ref. 2) (Fig. 3).

The supply of coarser clastics dramatically decreased in the latest Cambrian, and the Ordovician was characterized by slow deposition of black mudstones and cherts, (bulk of the Sydletsker Group¹ and formation B (ref. 2)) probably in a deeply submerged narrow rift. Thick resedimented conglomerates derived from the margin of the carbonate platform and the upper slope were deposited along the southern margins of the trough.

The transition from Ordovician to Silurian was marked by a new phase of turbidite sedimentation and a sequence several kilometres thick was deposited in the trough during the Silurian (formation C (ref. 2) and 'Un-named Silurian flysch formation'⁴). Silurian turbidites from eastern Peary Land to Hall Land (Fig. 1) were transported longitudinally from the east.

Contemporaneously, a variety of shallow marine limestones and dolomites with minor intercalations of terrigenous

clastics were deposited on the platform to the south (Figs 2 and 3).

During the Llandovery–Wenlock transition the platform margin collapsed. The position of the new border fault is not precisely known, but in the Peary Land area the onlapping turbidite sequence can be followed several hundred kilometres to the south⁶.

Silurian turbidite sedimentation was periodically interrupted by deposition of resedimented conglomerates derived from the upper slope and platform margin. The youngest preserved deposits of the North Greenland trough are turbidites and chert conglomerates.

Subsequently, orogenic deformation of the sequence is assumed to be of Devonian–early Carboniferous age³. Foldbelt strike is east–west parallel to the strike of the platform trough boundary. The southern margin of the foldbelt (Fig. 1) roughly coincides with this boundary and three major tectonic zones are recognizable in the Peary Land region. The platform carbonates are mainly undeformed whilst the trough clastics adjacent to the platform are folded. Northwards deformation increases and in Johannes V. Jensen Land the trough sediments are metamorphosed to amphibolite facies along the northern coast¹.

In the past, three independent stratigraphies have been put forward for each of these zones. To the south a comprehensive scheme covers the generally richly fossiliferous carbonates^{3,4,7–9}. The overlying Silurian turbidite sequence is lithologically monotonous and poorly fossiliferous and has not received the same detailed study. However, three formations can be recognized, comprising a basal mainly shaly unit of Llandovery to Wenlock age, a middle sandy turbidite unit of Wenlock age and an upper conglomerate unit of Wenlock age (Fig. 3).

The sequence of the foldbelt margin was subdivided into three formations². These include formation A, a sandy turbidite sequence capped by a major resedimented conglomerate unit. This is overlain by formation B, mainly a mudstone and chert sequence with wedges of resedimented conglomerate. Finally, formation C is the basal part of a major sandy turbidite unit. There are a few sporadic biostratigraphic dates from these sequences but there is no comprehensive zonation. Formation A is bioturbated in several horizons and it is underlain by units containing abundant trace fossils. The base of formation B contains graptolites of earliest Ordovician (early Canadian) age¹⁰. Consequently, formation A is probably Cambrian age throughout. Formation B has yielded graptolite faunas from the lower half indicating an early to middle Ordovician age. The graptolites *Climacograptus miserabilis* Elles & Wood and *Orthograptus* sp., resembling *O. quadrimucronatus* (Hall) indicating a late Ordovician (Cincinnatian) age, have been

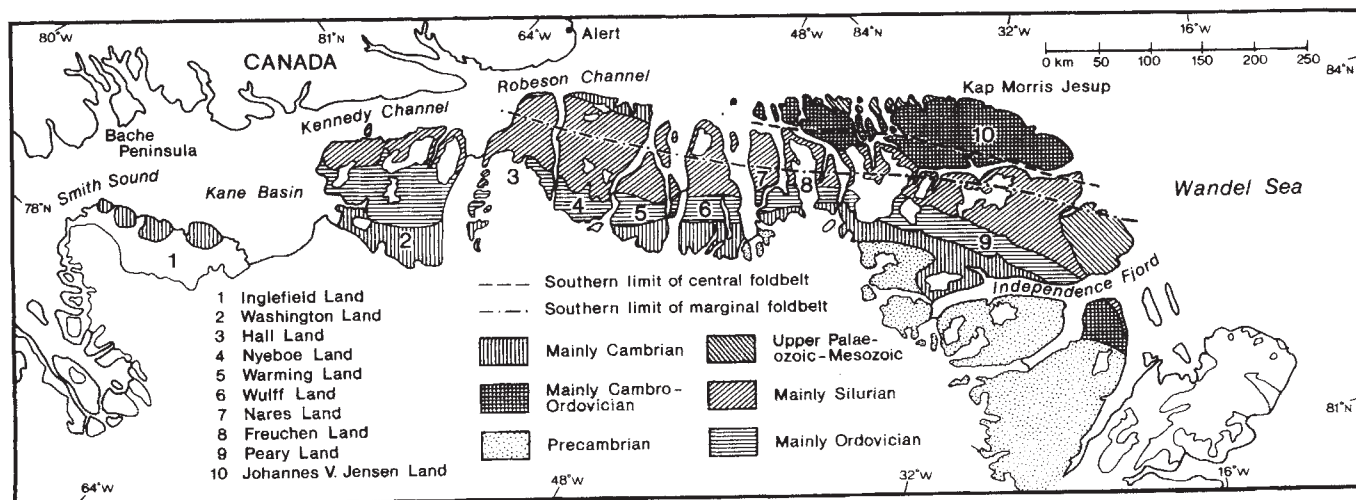


Fig. 1 Simplified geological map of North Greenland. Based on refs 3, 8 and unpublished observations. The southern limit of the central foldbelt corresponds to the Harder Fjord fault. The southern limit of the marginal foldbelt roughly corresponds to the Navarana Fjord fault.

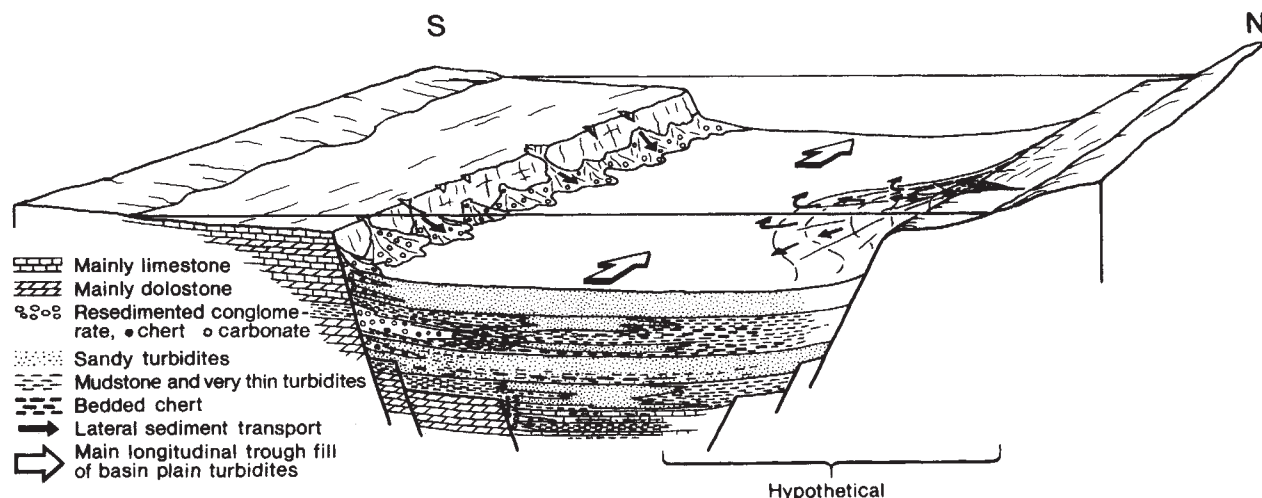


Fig. 2 Tectonic sedimentological setting of the early Palaeozoic rift in Peary Land, in late Silurian time.

obtained from the top of formation B. Formation boundaries are diachronous particularly between formations B and C, which fall within the Upper Ordovician to Lower Silurian range.

Until now, no age-diagnostic fossils have been obtained from the central part of the foldbelt, but unidentifiable graptolite fragments were reported from the upper part of the sequence¹. The first lithostratigraphic scheme for this area¹¹ was subsequently revised¹ and comprises three groups subdivided into informal formations (Fig. 3). A fourth group was recognized¹² in 1979. These comprise from below: an un-named quartzite group, and the Paradisfjeld, Polkorridoren and Sydgletscher Groups¹². Previously, correlation between the three independent stratigraphic schemes of the central foldbelt, foldbelt margin and unfolded carbonates and clastics to the south was speculative. This was due to lack of faunas and to lateral facies changes.

Two formations were recognized in the Polkorridoren Group¹, of which the upper characteristic turbidite unit was tentatively correlated with formation C of the foldbelt margin and the Silurian turbidites of the unfolded areas to the south³. This correlation was based primarily on lithological resemblance. By implication the overlying Sydgletscher Group was

assumed to be of late Silurian to Devonian age^{2,3}. This group which is more than 1 km thick is divided into five formations^{1,12}. An assemblage of graptolites was found 142 m above the base of the penultimate formation, termed upper Sydgletscher shales (3d)¹. These graptolites, the first age-diagnostic fossils derived from the central part of the foldbelt (Fig. 4), provide a correlation link with areas to the south, considerably revising earlier schemes.

The graptolites appear as flattened carbon films with inconspicuous thecal structures and some are tectonically compressed. The material comprises about 10 incomplete biserial and 15 monoserial fragments. Three specimens can be referred to *Climacograptus rectangularis* (M'Coy)? and four monoserial fragments to *Atavograptus* aff. *atavus* Jones (Fig. 4). *C. rectangularis* has generally been reported from the *atavus* to *cyphus* Zones of the Lower Llandovery¹³ and recently at least up to the *triangulatus* Zone¹⁴. *A. atavus sensu lato* has been reported from the *atavus* to *magnus* Zones¹³. However, *A. aff. atavus* is very similar to specimens described from the *atavus* Zone of Bornholm¹⁵ and the sample is probably referable to a stratigraphical level around the *atavus*-*cyphus* Zones in the Lower Llandovery.

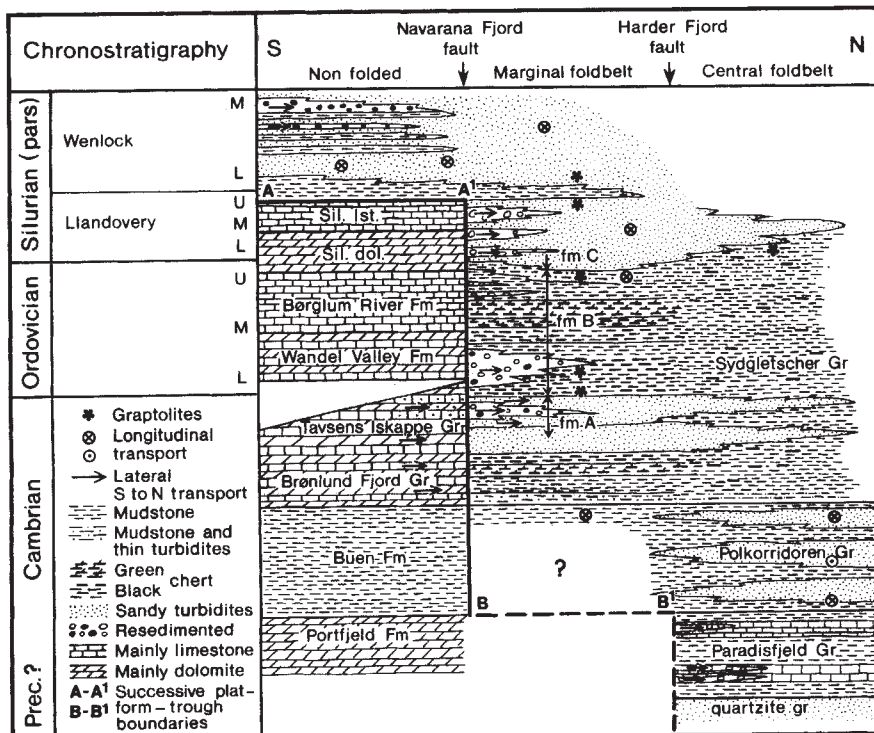


Fig. 3 Stratigraphy and facies distribution of a north-south transect through the Lower Palaeozoic carbonate platform-clastic trough system of Peary land. Based on refs 1, 4, 22 and unpublished observations.

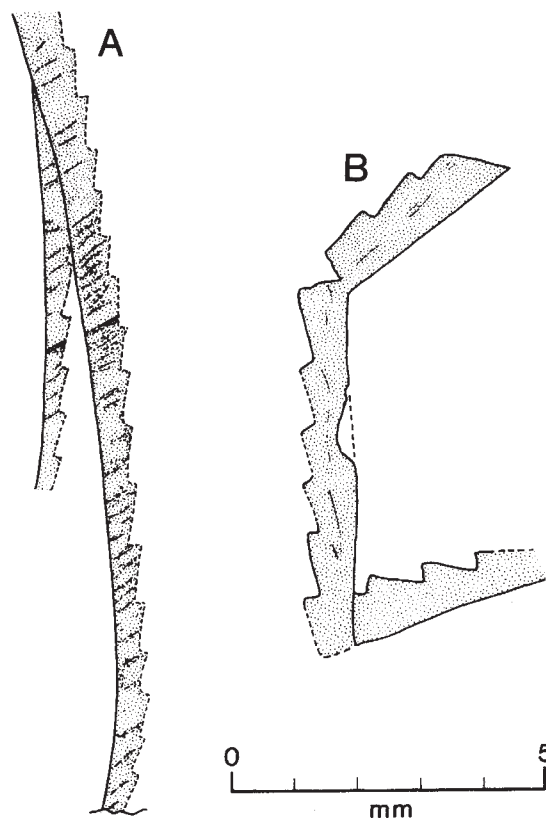


Fig. 4 First age diagnostic fossils from the central North Greenland foldbelt. *a*, *Atavograptus* aff. *atavus* (Jones). Two proximal fragments. *b*, *Atavograptus* aff. *atavus*. One distal fragment. Both specimens are from the upper Sydgletscher shales. MGUH 15077 and MGUH 15078. GGU sample 230273.

This indicates the bulk of the Sydgletscher Group is of pre-Lower Llandovery, *atavus* Zone age, and thus mainly of Ordovician age and that it is broadly contemporaneous with formation B of the foldbelt margin (Fig. 3). The correlation between the Sydgletscher Group and formation B is strongly supported by lithological evidence as both units are characterized by black cherts and mudstones. The graptolite-bearing formation is overlain by a several hundred metres thick formation consisting of extremely thick turbidite beds. Individual beds are structureless and reach thicknesses of 30 m suggesting a rapid depositional rate for the formation.

This formation is the youngest rock unit of the central foldbelt and it is unlikely to be much younger than early Llandovery.

The three stratigraphic schemes used to describe the trough sediments of the central foldbelt, foldbelt margin and unfolded area can now be combined into a single stratigraphic scheme (Fig. 3). The scheme which is informal, shows the main facies types and their spatial distribution. A formal comprehensive stratigraphy will be published elsewhere.

In spite of the new biostratigraphic data, it is still difficult to date precisely the orogenesis. The youngest folded rocks within the central foldbelt are of Llandovery or only slightly younger age. Towards the south and west in the marginal parts of the foldbelt, the top part of the Silurian turbidite sequence yields graptolites of Ludlow (late Silurian) age (our unpublished observations). The presence of earliest Devonian has been suggested for a single locality in western Hall Land on the basis of one specimen of the graptolite *Monograptus* cf. *aequalis*¹⁶. It was recently reidentified by Jaeger (personal communication) as *Monograptus* cf. *transgrediens* Perner and probably of Pridoli (latest Silurian) age. The graptolite bed was overlain by strata containing alleged Devonian vertebrates¹⁷ assigned a Devonian age partly on the basis of the underlying graptolites. There is thus no evidence of Devonian rocks in North Greenland.

In eastern Peary Land the folded sequences are overlain with angular unconformity by Upper Pennsylvanian and younger rocks^{3,18-20}. Reliable isotopic ages have not yet been obtained. Therefore, the diastrophism can only be loosely dated as Devonian to early Pennsylvanian. Comparison with Arctic Canada fails to yield more definitive evidence regarding time of diastrophism. In Ellesmere Island the most extensive and intensive deformation in the history of the Innuian mobile belt commences not later than middle Devonian and was completed before the late early Pennsylvanian²¹.

We thank P. R. Dawes, J. D. Friderichsen, N. Henriksen, A. K. Higgins and J. S. Peel for criticisms. This work is published with the permission of the director of the Geological Survey of Greenland.

Received 5 May; accepted 26 June 1980.

1. Dawes, P. R. & Soper, N. J. *Mem. Am. Ass. petrol. Geol.* **19**, 117-134 (1973).
2. Dawes, P. R. & Soper, N. J. *Rapp. Grönl. geol. Unders.* **93**, 1-40 (1979).
3. Dawes, P. R. in *Geology of Greenland* (eds Escher, A. & Watt, W. S.), 248-303 (Geological Survey of Greenland, Copenhagen, 1976).
4. Christie, R. L. & Peel, J. S. *Rapp. Grönl. geol. Unders.* **82**, 1-48 (1977).
5. Trettin, H. P. *et al. Geol. Ass. Can. Spec. Pap.* **11**, 83-179 (1972).
6. Peel, J. S. *Rapp. Grönl. geol. Unders.* (in the press).
7. Troelsen, J. C. *Meddr Grønland* **149**, 1-29 (1949).
8. Dawes, P. R. *Bull. geol. Soc. Denmark* **20**, 197-239 (1971).
9. Jepsen, H. F. *Meddr Grønland* **192**, 1-42 (1971).
10. Bjerreskov, M. & Poulsen, V. *Rapp. Grönl. geol. Unders.* **55**, 10-14 (1973).
11. Fränkl, E. *Meddr Grønland* **103**, 1-61 (1955).
12. Soper, N. J. *et al. Rapp. Grönl. geol. Unders.* (in the press).
13. Hutt, J. *Palaeontogr. Soc. (Monogr.)*, pt. 2, 1-59 (1975).
14. Lenz, A. C. *Acta Palaeont. Pol.* **24**, 137-153 (1979).
15. Bjerreskov, M. *Fossils and Strata* **8**, 1-94 (1975).
16. Berry, W. B. N. *et al. Rapp. Grönl. geol. Unders.* **65**, 11-13 (1974).
17. Bendix-Almgreen, S. E. & Peel, J. S. *Rapp. Grönl. geol. Unders.* **65**, 13-16 (1974).
18. Troelsen, J. C. in *Geology* (eds Winther, P. C. *et al.*) 6-8 (1950).
19. Petryk, A. A. *Rapp. Grönl. geol. Unders.* **85**, 16-21 (1977).
20. Håkansson, E. *Rapp. Grönl. geol. Unders.* **88**, 73-83 (1979).
21. Trettin, H. P. & Balkwill, H. R. *Can. J. Earth Sci.* **16**, 749-769 (1979).
22. Ineson, J. & Peel, J. S. *Rapp. Grönl. geol. Unders.* (in the press).

Structure of supernumerary limbs

M. Maden

The National Institute for Medical Research, Mill Hill, London NW7 1AA, UK

The generation of supernumerary limbs by rotation of amphibian limb buds or regeneration blastemas is a phenomenon well known to embryologists^{1,2}. The recent revival of interest in this aspect of limb morphogenesis has been largely due to the stimulus provided by the polar coordinate model for control of limb outgrowth and pattern³ and subsequent efforts to test its validity. The geometrical rules postulated in the polar coordinate model make precise predictions about the handedness of supernumerary limbs. Following antero-posterior (AP) or dorso-ventral (DV) inversions, they should be of the same handedness as the stump, whether single or in pairs, but after inversion of both axes (APDV) pairs should be of opposite handedness³. It is surprising, therefore, that the structure of supernumerary limbs has never been studied in sufficient detail to report whether they really are normal limbs, for only the skeletal elements are examined in cleared whole mounts (see Fig. 1). This only allows an assessment of the AP axis, and to determine the dorsal and ventral surfaces (and hence the handedness), the curvature of the digits is noted. I therefore decided to study the muscle patterns of all types of supernumeraries, and report here that, contrary to expectation, the majority of APDV supernumeraries are not normal limbs but are mirror-image duplications in the dorso-ventral axis, that is, either double-dorsal or double-ventral. This presents problems for contemporary models of pattern formation in the limb, and issues a warning to those who use supernumerary limbs as an experimental tool to investigate other areas of developmental biology.

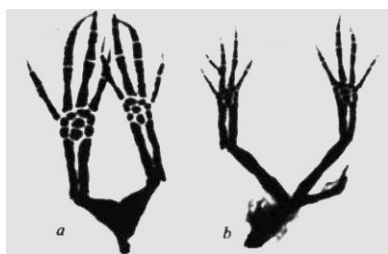


Fig. 1 *a*, Supernumerary limb generated by grafting a left blastema to a right stump. On the left is the right graft, on the right the perfect right-hand supernumerary limb which appeared at the posterior pole of the stump and bifurcates in the stylopodium. *b*, Supernumerary limb generated by rotating a right blastema 180° and replacing it on the same stump. On the right is the graft, on the left the supernumerary which appeared at the dorsal pole of the stump and bifurcated in the stylopodium. The similarity between this supernumerary and that in *a* is striking and it would normally have been identified as a left limb. In fact, on sectioning it was discovered to be double-ventral in structure (see Figs 2c, 3e). Victoria blue staining.

The experiments were carried out on young *Ambystoma mexicanum*, 80–140 mm long, whose limbs had been amputated through the mid-stylopodium level. At the late cone/early palette stage, blastemas were cut off and either placed on the contralateral stump (AP or DV inversions) or replaced upside down on the same stump (APDV rotations). Such grafted blastemas continue their own development into normal limbs as well as participating in the generation of supernumerary limbs at the graft–stump boundary. When they had fully developed, the muscle patterns of 20 of these grafts were examined and they revealed a completely normal anatomy⁴. The most characteristic features of normal limbs are to be found at the carpal and digit level, where the dorsal extensors have separated into four crescent-shaped muscles, the extensores digitorum breves (Fig. 2a). The flexors, comprised of about 16 separate muscles⁴, form a large mass of tissue on the ventral surface of the hand which in the digits, splits into four semi-circles of muscle (Fig. 2a). A three-dimensional reconstruction of the normal muscle pattern (Fig. 3a) shows that the sparser dorsal muscles terminate in the metacarpals (tendons travel distally to the tips of the digits) whereas the ventral muscles themselves extend right to the last phalanx.

The first series consisted of 11 limbs whose AP axis had been reversed, yielding a total of 12 supernumeraries. Ten limbs developed single supernumeraries, all at the posterior side of the host limb (Fig. 1a), and one a double supernumerary (Table 1). It is interesting that this striking bias towards the production of

supernumeraries at the posterior pole after upper arm inversions is also apparent from other studies^{5,6}, but it disappears when blastemas from the lower arm are inverted. All the supernumeraries produced in this series were perfect limbs whose digits curved ventrally and from an examination of the skeletal structure (Fig. 1a) would be considered normal limbs. Analysis of the muscle patterns of each of these 12 supernumeraries confirmed that they were indeed completely normal (Fig. 3b, Table 1) and that all were of stump handedness.

The second series comprised seven limbs whose DV axis had been reversed. This yielded a total of 10 supernumeraries, 4 single and 3 double (Table 1). These were more equally distributed between the two poles than in the previous series—55% dorsal, 45% ventral. One was slightly displaced in the antero-dorsal quadrant. Apart from their position, these supernumeraries were in every respect identical to those in the first series and an analysis of the muscle pattern confirmed that they were completely normal limbs (Fig. 3c, Table 1). Again, they were all of stump handedness.

The third series comprised 12 blastemas rotated 180° (APDV). These were a group of limbs from a previous series⁷ and were selected for analysis here because of a good separation of graft and supernumerary(ies). For instance, the example shown in Fig. 1b bifurcates in the humerus. Eight limbs had single supernumeraries and four had double supernumeraries, making 16 in all, whose position of origin was random (Table 1). These all looked normal when stained with Victoria blue (Fig. 1b) although in some cases their digits did not curve. An analysis of their muscle patterns revealed that 5 of the 16 (Table 1) had at the carpal level separate crescent-shaped extensores digitorum breves both dorsally and ventrally (Fig. 2b), terminating at the metacarpals. These limbs had no ventral flexor muscles extending to the tips of the digits and were thus identified as double-dorsal limbs. The three-dimensional reconstructions supported this conclusion (Fig. 3d). A further six supernumeraries had an uncharacteristically excessive amount of muscle on their dorsal surfaces which did not separate into small crescent-shaped elements (Fig. 2c). Instead of the dorsal muscles terminating in the metacarpals, they continued to the end of the digits and, together with the normal ventral muscles, formed thick circles of muscle surrounding each digit (Fig. 2c). These limbs were thus identified as double-ventral (Fig. 3e, Table 1). The third category of APDV supernumeraries contained five limbs which were completely normal (Fig. 2d, Table 1).

Thus, the AP and DV supernumeraries examined here were normal in every respect and were of stump handedness as had been previously supposed. In contrast, the APDV supernumeraries had an equal chance of being double-dorsal, double-

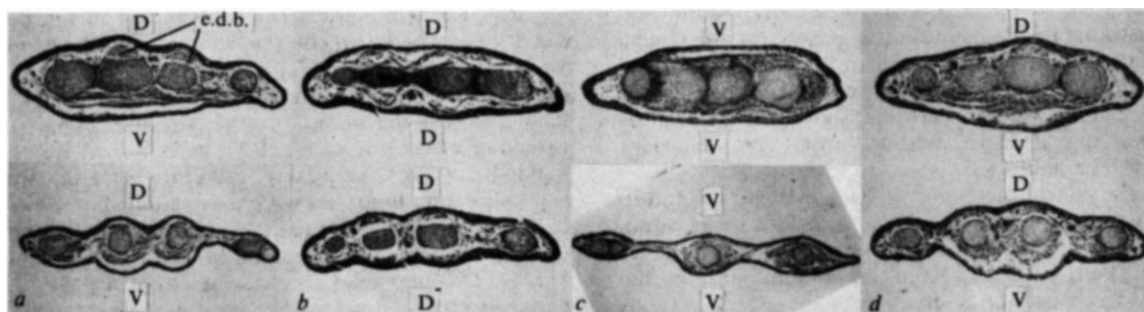


Fig. 2 Sections through the carpal/metacarpal level (upper sections) and distal metacarpal level (lower sections) of various types of limb. Stained with haematoxylin and eosin. The dorsal (D) and ventral (V) surfaces are marked. *a*, A limb which developed from a grafted blastema. This shows completely normal muscle patterns⁴. In the upper section note that the first (left) and fourth (right) digit extensors have already terminated, leaving only those of the second and third digits (e.d.b.). These are crescent-shaped and sit on the flat tops of the basal carpals. The ventral flexor muscles extend right across the hand and are composed of about 16 separate muscles⁴. The lower section through the digits shows the ventral muscles separated into large semi-circles. $\times 175$. *b*, An APDV supernumerary. Clearly, the muscle patterns have mirror-image symmetry, as small, crescent-shaped muscles appear at the top and bottom of the limb. This is a double-dorsal limb, the three-dimensional reconstruction of which is shown in Fig. 3d. Note also that, particularly in the lower section, the cartilages seem to be mirror-images—both surfaces of the middle two elements are flat-topped. $\times 168$. *c*, Another type of APDV supernumerary. As in *b* the muscles are mirror-images and in the digits they form complete circles of muscle around the phalanges. This is a double-ventral limb, the three-dimensional reconstruction of which is shown in Fig. 3e. Again, particularly in the upper section, the carpals appear mirror-images. $\times 120$. *d*, Another type of APDV supernumerary. This, in contrast to the previous two, was a completely normal limb (compare with *a*). $\times 190$.

Table 1 The number, position (relative to stump axes) and structure of supernumeraries generated by AP, DV or APDV rotation of blastemas

No.	AP supernumeraries				DV supernumeraries				APDV supernumeraries			
	Supernumerary 1		Supernumerary 2		Supernumerary 1		Supernumerary 2		Supernumerary 1		Supernumerary 2	
	Position	Structure	Position	Structure	Position	Structure	Position	Structure	Position	Structure	Position	Structure
1	P	Normal	—	—	V	Normal	—	—	PD	Double D	—	—
2	P	Normal	—	—	AD	Normal	—	—	AV	Double D	—	—
3	P	Normal	—	—	D	Normal	—	—	D	Normal	—	—
4	P	Normal	—	—	D	Normal	—	—	D	Normal	—	—
5	A	Normal	P	Normal	D	Normal	V	Normal	D	Double D	—	—
6	P	Normal	—	—	D	Normal	V	Normal	D	Double V	—	—
7	P	Normal	—	—	D	Normal	V	Normal	AD	Double D	—	—
8	P	Normal	—	—	—	—	—	—	V	Double V	—	—
9	P	Normal	—	—	—	—	—	—	AV	Double D	P	Double V
10	P	Normal	—	—	—	—	—	—	D	Normal	V	Normal
11	P	Normal	—	—	—	—	—	—	A	Double V	V	Double V
12	—	—	—	—	—	—	—	—	P	Normal	D	Double V

P, posterior; A, anterior; D, dorsal; V, ventral; AD, antero-dorsal quadrant; AV, antero-ventral quadrant; PD, postero-dorsal quadrant.

ventral or normal. As emphasized above, only well separated, four-digit supernumeraries were selected for study and the typical hypomorphic structures produced in any series of rotations—spikes, fused supernumeraries or ones with fewer digits than normal—were ignored. These are clearly abnormal in the AP axis and perhaps also in the DV axis. Furthermore, it may well be that other classes of APDV supernumeraries will be found, those appearing at frequencies of less than 1 in 16, and this possibility is now being investigated. Nevertheless, the generation of mirror-image limbs has two significant implications for theories of pattern formation. First, it adds to the growing body of results which present severe difficulties for the polar coordinate model and is contrary to the prediction that pairs of APDV supernumeraries should be of opposite handedness. The concept of handedness is meaningless when applied to mirror-image limbs, and pairs were not even of opposite mirroring (Table 1). The only other model which attempts to describe the generation of supernumerary limbs in such detail is the serial threshold model⁶, which ascribes the formation of mirror-image limbs to different rates of intercalation in each of the three axes. However, theories based

solely on local cell-cell interactions cannot, at their present stage of development, provide a satisfactory explanation for the appearance of APDV supernumeraries because, whatever their structure, there will always be large areas of gross discrepancy either between the supernumerary and the stump or between the supernumerary and the blastema. This problem is not solved by the discovery of mirror-image limbs because a double-dorsal limb at, say, the dorsal pole of the stump will abut directly on to the ventral pole of the rotated blastema. This should, according to local intercalation rules, generate more supernumeraries *ad infinitum*. Clearly, such simplistic descriptions cannot explain these phenomena and new models must be devised which incorporate more global field constraints.

These present results also suggest that deterministic models be replaced by those with a more probabilistic nature, however unsatisfying this may be. The polar coordinate model is a highly deterministic model in which the cells can only show one particular behaviour following one particular experimental interference. The result is a strict series of predictions. This is clearly not the case for APDV supernumeraries; each class of structure—double-dorsal, double-ventral or normal—had an equal chance of appearing. This probabilistic concept has been considered previously in connection with the induction of supernumeraries at angles of rotation other than 180° (ref. 7). Further work has revealed a frequency distribution in which the angle of rotation is directly related to the likelihood of supernumerary induction⁸.

It is to be hoped that this finding of mirror-image supernumeraries is a universal phenomenon in both regeneration and development. In support of this, I have carried out 180° rotations of developing limb buds in *Rana temporaria* and nearly all of the supernumeraries so generated were double posterior in structure. This may indicate that the plane in which APDV supernumeraries are duplicated is species dependent, but that the general principle remains the same. If this is so, it has implications for the other developing system commonly studied, the chick limb bud. It means that a graft of the zone of polarizing activity (ZPA) in normal dorso-ventral orientation (equivalent to an AP inversion) should generate a different type of reduplicate from a ZPA graft with reversed dorso-ventral orientation (equivalent to an APDV rotation). How current models of ZPA activity can cope with this situation remains to be seen.

I thank Katriye Mustafa for technical assistance in all aspects of this work.

Received 12 March; accepted 20 June 1980.

1. Abeloos, M. & Lecamp, M. C. *r. hebdo. Séanc. Acad. Sci., Paris* **192**, 639–641 (1931).
2. Harrison, R. G. *J. exp. Zool.* **32**, 1–136 (1921).
3. French, V., Bryant, P. J. & Bryant, S. V. *Science* **193**, 969–981 (1976).
4. Grim, M. & Carlson, B. M. *Z. Anat. Entwickl.-Gesch.* **145**, 137–184 (1974).
5. Tank, P. W. *Dev. Biol.* **62**, 143–161 (1978).
6. Wallace, H. & Watson, A. *J. Embryol. exp. Morph.* **49**, 243–258 (1979).
7. Maden, M. & Turner, N. *Nature* **273**, 232–235 (1978).
8. Slack, J. M. W. *J. theor. Biol.* **82**, 105–140 (1980).
9. Turner, N. & Maden, M. (submitted).

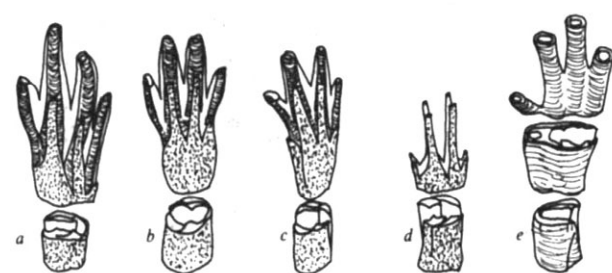


Fig. 3 Three-dimensional reconstructions of muscle patterns from serial sections of various limbs. Every 10th section was drawn with a camera lucida. These drawings were then projected 1 cm apart on to another piece of paper and the surface configurations of the muscles were then drawn. Each reconstruction drawn above is to the same scale to show the variation in proximo-distal extent of muscles. They are all viewed from the dorsal (extensor) surface with flexors underneath and all have a break in the wrist to help visualize their internal structure. *a*, Grafted limb. The blastema, grafted at the late cone/early palette stage, develops into a normal limb despite having at least one axis inverted relative to the stump. The dorsal extensor muscles (dotted) terminate at the carpal/metacarpal level whereas the ventral flexor muscles (cross-hatched) continue to the ends of the digits. This is the normal muscle pattern in the axolotl forelimb⁴. *b*, An AP supernumerary limb. This shows the same features as *a*—short extensors (dotted), long flexors (cross-hatched). *c*, A DV supernumerary limb showing normal muscle patterns as in *a* and *b*. *d*, An APDV supernumerary. In this case, the ventral muscles (clear) do not extend to the tips of the digits, but form a mirror-image pattern of the dorsal muscles (dotted) and terminate at the carpal/metacarpal level. This is a double-dorsal limb, the histology of which is shown in Fig. 2*b*. *e*, An APDV supernumerary. Here both dorsal (cross-hatched) and ventral muscles are much more profuse; they continue to the tips of the digits and form complete circles of muscle. This is a double-ventral limb, the histology of which is shown in Fig. 2*c*.

A one-dimensional pattern in the cellular slime mould *Polysphondylium pallidum*

F. W. Spiegel & E. C. Cox

Department of Biology, Princeton University, Princeton, New Jersey 08544

The cellular slime mould *Polysphondylium pallidum* has a fruiting body with whorls of branches spaced at regular intervals along the stalk (Fig. 1). The spacing between successive whorls is quite regular and under genetic control¹. Because stalk cell number per interval correlates strongly with interval length, we have studied the influence of ploidy on the distance between whorls to determine if cell counting could account for the spacing. It is known that diploid spores and amoebae of *Dictyostellium discoideum* and *Polysphondylium violaceum* are roughly twice the size of their haploid parents^{2,3}, and we find that stalk cells in *P. pallidum* follow the same pattern. Our results show that the number of stalk cells between successive whorls in haploids is approximately twice the number in diploids, while the spacing remains the same. This clearly indicates that spacing is not achieved by a cell counting mechanism. Instead, it appears to depend on distance.

Cells were grown and diploids were constructed by previously published methods³ using genetically marked parent strains. In every case diploid clones were shown to segregate the appropriate recessive genetic markers. Two- to three-day-old fruiting bodies cultured on water agar in the presence of activated charcoal were either photographed directly (Fig. 1) or trapped under a cover slip for photography at higher magnification (Fig. 2). Figure 1 compares developing and mature diploid fruiting

bodies to the two haploid parents. It is quite clear that whorl spacing in diploids is very similar to that in haploids. Quantitative measurements of large numbers of sorocarps bear out this impression: spacing in five independently isolated diploids is indistinguishable from the haploid parents (Table 1).

The most striking feature of whorl spacing in diploid sorocarps is that spacing does not depend on a cell counting mechanism. This conclusion follows from the results shown in Fig. 2 and Table 1. It is qualitatively evident from Fig. 2 that the number of cells between whorls is larger in the two haploid parents than in diploid progeny, even though the distance between whorls is very similar. This impression is borne out by the detailed quantitative measurements summarized in Table 1.

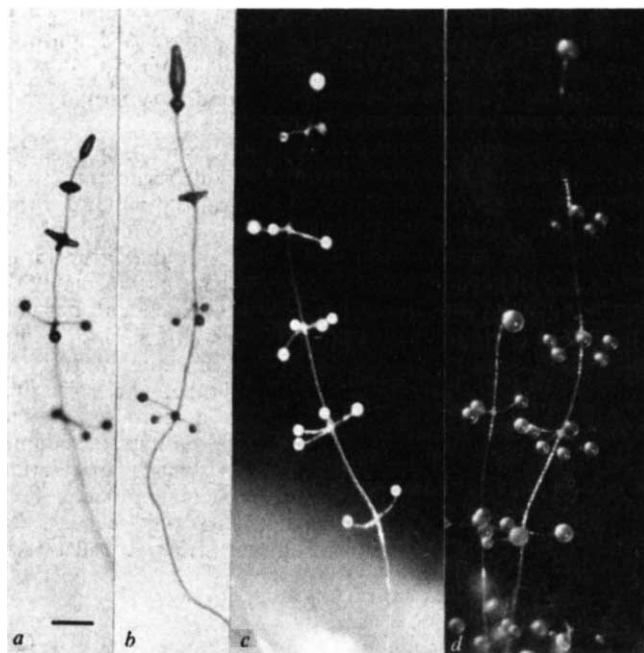


Fig. 1 Comparison of haploid and diploid *Polysphondylium pallidum* fruiting bodies. *a*, Culminating sorogen of haploid PN519; *b*, culminating sorogen of diploid 95a; *c*, mature sorocarp of haploid PN525; *d*, mature sorocarp of diploid 95a. The larger sori in *d* are the result of absorption of water in humid culture conditions. Scale bar, 100 μ m.

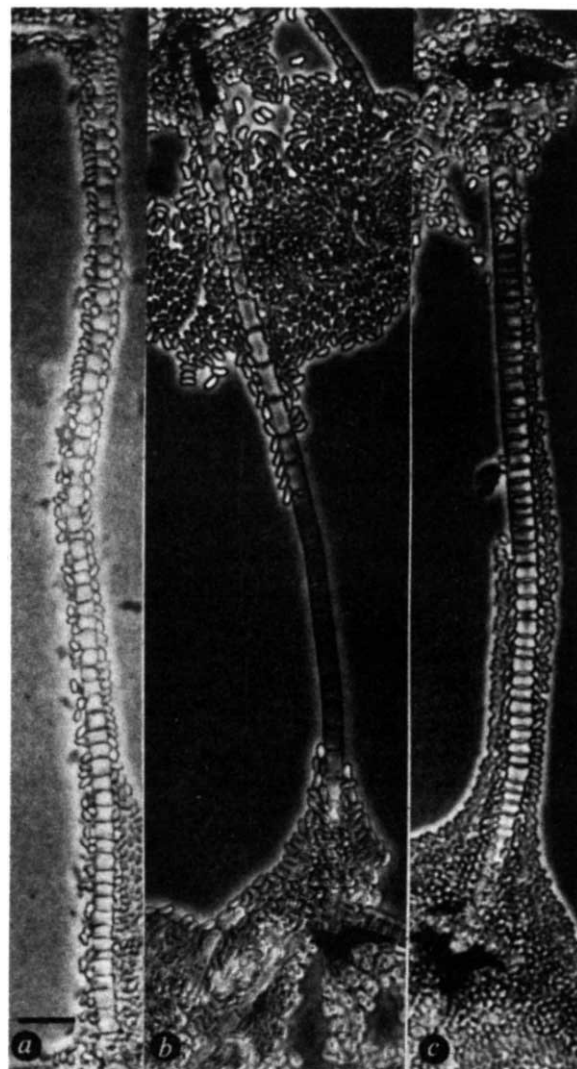


Fig. 2 Comparison of stalk cell size in *Polysphondylium pallidum* haploid (*a, c*) and diploid (*b*) fruiting bodies. Stalk intervals of haploid PN525 (*a*), diploid 95a (*b*), and haploid PN519 (*c*). Note the larger spores of the diploid. Scale bar, 15 μ m.

The average number of cells per interval in the haploid strains PN519 and PN525 is almost exactly twice the number in the diploid isolates 95a to 95e, while the cell length is approximately one half. The size and shape of diploid spores, shown clustered about the stalks in Fig. 2, bears the same relationship to haploid spore shape in *P. pallidum* as it does in *D. discoideum* and *P. violaceum*^{2,3}.

These results thus eliminate one simple hypothesis that is *a priori* appealing, and suggest that whorl spacing depends on

some mechanism other than cell counting. They are reminiscent of the effect of increasing ploidy on cell number in amphibians, where animals maintain their size and shape by using fewer, larger cells⁴. *P. pallidum* is unlike many higher plants, however, which tend to become proportionately larger following an increase in ploidy^{5,6}, and unlike bristle spacing along *Drosophila melanogaster* basitarsi, where spacing mechanisms seem to depend on the number of cells in normal diploid and intersex triploids⁷.

Table 1 Interval length between whorls and cell number per interval in haploid and diploid strains of *P. pallidum*

Strain	Ploidy	Average interval length (μm)	Average cell number/interval	Average cell length/interval (μm)
PN519	<i>n</i>	374 (110)	45.7 (18.5)	8.7 (2.2)
PN525	<i>n</i>	354 (79)	40.4 (14.0)	9.5 (2.7)
95a	2 <i>n</i>	350 (99)	18.4 (8.1)	20.7 (5.8)
95b	2 <i>n</i>	319 (100)	20.5 (8.0)	16.7 (4.7)
95c	2 <i>n</i>	349 (87)	21.4 (9.7)	18.1 (5.3)
95d	2 <i>n</i>	358 (97)	20.5 (7.5)	18.8 (5.3)
95e	2 <i>n</i>	304 (91)	18.8 (8.3)	17.4 (4.8)

Comparison of interval length between whorls, stalk cell number per interval, and stalk cell length between haploid and diploid strains of *P. pallidum*. A minimum of 100 intervals on a minimum of 20 sorocarps was counted for each strain. Diploids 95a-e were constructed from the parents PN519 and PN525³. Values shown in parentheses represent $\pm$ s.d.

The developmental history of each whorl in *P. pallidum* is complex. Stalk cells are laid down apically as the sorogen ascends the elongating stalk. At regular intervals masses containing variable numbers of cells pinch off basally from the ascending sorogen (this may be seen clearly in Fig. 1a). These cells differentiate into whorls with varying numbers of arms, each arm bearing a terminal sorus of spores^{8,9}. The timing of these events clearly determines the spacing frequency. We have determined¹ that the control of spacing in *P. pallidum* is as well defined as heterocyst spacing in the prokaryote *Anabaena cylindrica*¹⁰, even though the distances over which the spacing mechanism must operate are roughly an order of magnitude larger in *P. pallidum*, because spacing mutants can be isolated in *P. pallidum* (our unpublished data). Since a facile genetic system is available¹¹, and because the spacing may be altered by physiological manipulation^{9,12}, it may be possible to analyse in some detail the genetic and physical basis for whorl spacing in this simple eukaryote.

We thank David Francis for supplying the genetic stocks used in this study. This work was supported in part by a grant to the Biology Department by the Whitehall Foundation and by USPHS grants AI09392 and CA09167.

Received 28 March; accepted 11 July 1980.

1. Spiegel, F. W. & Cox, E. C. (in preparation).
2. Sussman, M. & Sussman, R. R. *J. gen. Microbiol.* **28**, 417-429 (1962).
3. Warren, A. J., Warren, W. D. & Cox, E. C. *Genetics* **83**, 25-47 (1976).
4. Fankhauser, G. in *Analysis of Development* (eds Willier, B. H., Weiss, A. & Hamburger, V.) 126-150 (Saunders, Philadelphia, 1955).
5. Darlington, C. D. *Chromosome Botany and the Origins of Cultivated Plants*, 3rd edn (Allen and Unwin, London, 1973).
6. Stebbins, G. L. *Chromosomal Evolution in Higher Plants* (Addison-Wesley, Reading, Massachusetts, 1971).
7. Held, L. I. Jr *Wilhelm Roux Arch. EntwMech. Org.* **187**, 105-127 (1979).
8. Bonner, J. T. *The Cellular Slime Molds* (Princeton University Press, 1967).
9. Harper, R. A. *Bull. Torrey Bot. Club* **59**, 49-84 (1932).
10. Wilcox, M. *Nature* **228**, 686-687 (1970).
11. Francis, D. W. *Differentiation* **15**, 187-193 (1970).
12. Raper, K. B. in *The Fungi* (eds Ainsworth, G. C., Sparrow, F. K. & Sussman, A. S.) 9-36 (Academic, New York, 1973).

Induction of abnormal development and differentiation in cultured mammary glands by cyclic adenine nucleotide and prostaglandins

Frederick V. Schaefer, R. Philip Custer & Sam Sorof

Institute for Cancer Research, Fox Chase Cancer Center, Philadelphia, Pennsylvania 19111

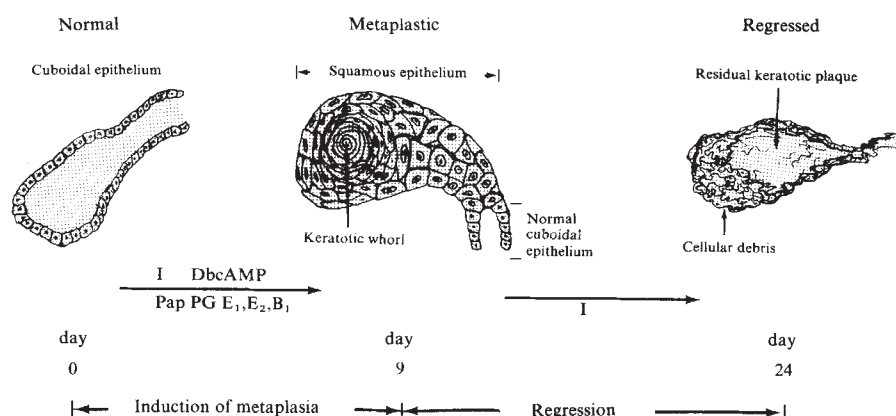
Squamous (epidermoid) metaplasia is a common abnormality of development and differentiation, in which various epithelia are replaced by cells resembling those of the epidermis. The metaplastic cells differentiate into flattened cells that frequently form keratotic flakes and whorls. The cells may regress, remain benignly squamous, or progress to epidermoid carcinoma¹. Squamous metaplasias have previously been brought about by mechanical injury², vitamin A deficiency^{3,4} and chemical carcinogens⁵⁻⁷. In search of the physiological mediators of abnormal development and differentiation in oncogenesis, we have now discovered that the combination of dibutyryl cyclic AMP, prostaglandins (PGs) E₁, E₂ and B₁, and papaverine (Pap) induces squamous metaplasia with keratin production in the epithelium of cultured mammary glands. The three inducers may act synergistically to elevate the level of intracellular cyclic adenine nucleotide. Removal of these inducers causes the regression of the metaplastic cells. Retinoid does not block the induction of this squamous metaplasia, even though retinoid has previously prevented and reversed the squamous metaplasias caused by vitamin A deficiency or chemical carcinogens in other organs^{4,8}. The findings suggest that cyclic adenine nucleotide and the PGs may mediate the squamous metaplasias present in many abnormal, transformed and malignant tissues, and possibly also the normal development and differentiation of epithelia into keratin-producing, stratified squamous cells, as in skin.

The mouse mammary gland can now be manipulated to undergo three courses of development and differentiation *in vitro*. First, normal development and differentiation of mouse mammary glands can be induced in culture^{7,9,10}. The glands of hormonally primed female BALB/c mice, 3 to 4 weeks old, consist mainly of ducts with numerous terminal buds. When cultured in serum-free medium containing insulin (I), prolactin (P), aldosterone (A) and hydrocortisone (H), the glands develop lobuloalveoli^{7,9,10} and produce casein¹¹. Removal of all hormones except I (regression medium) brings about alveolar involution^{7,10}. Such mammary glands exhibit no metaplasia (series 1 and 2, Table 1). Subsequent incubation with epidermal growth factor and IPAH induces a second lobuloalveolar development, which can again be involuted in the regression medium¹². All these processes resemble those that normally occur *in vivo*.

Second, the epithelial cells of the cultured mammary glands can also be transformed by chemical carcinogens, as evidenced by the production of nodule-like alveolar lesions that have escaped the hormonal controls of alveolar development^{7,13-15}. The transformed glands are dysplastic, metaplastic^{7,15} and tumorigenic¹⁴. In contrast, chemically analogous noncarcinogens have little or no transforming activity^{7,16}. Retinoid can prevent, suppress and apparently reverse the transformations caused by procarcinogens^{15,17}, which require metabolic activation to become reactive, but not the transformations brought about by many activated carcinogens¹⁷.

A third course of development and differentiation is reported here. The combination of three agents (dibutyryl cyclic AMP, specific PGs and Pap) induces an extensive stratified squamous cell metaplasia with keratin production in cultured mammary glands (Fig. 1). Fresh whole mammary glands of hormonally

Fig. 1 Abnormal development, differentiation and involution in mouse mammary glands in culture: induction of squamous metaplasia and keratin production by cyclic adenine nucleotide, PGs and Pap. The serum-free whole organ culture system has been described^{7,9,10}. Female BALB/c mice, 3–4 weeks old, were primed by subcutaneous injections of β -oestradiol (1 μ g) and progesterone (1 mg) for 9 consecutive days. On the day following the last injection (day 0), mice were killed and their entire second thoracic mammary glands were excised and floated in medium on Dacron rafts in 95% O₂ and 5% CO₂ atmosphere. Serum-free keratinizing medium consisted of Waymouth MB752/1 medium (1 ml per gland), 350 μ g ml⁻¹ L-glutamine, 35 μ g ml⁻¹ penicillin G, and the supplements, dibutyl cyclic AMP (DbcAMP) (10⁻⁴ M), Pap (10⁻⁶ M), (both Sigma), and I (5 μ g ml⁻¹) (Calbiochem). 'Enhancing medium' consisted of the keratinizing medium with additions of PGE₁, PGE₂ and PGB₁ (each 5 μ g ml⁻¹) (Upjohn). On day 9, the mammary glands exhibited extensive squamous metaplasia and deposits of keratin. Maintenance of the glands for an additional 15 days (days 9–24) in 'regression medium', containing I (5 μ g ml⁻¹) as the only supplement, caused the metaplastic foci to involute. Media were renewed every 2 or 3 days as required. Glands were coded, fixed, sectioned at 5 μ m, and stained with haematoxylin and eosin for histological analysis, and by Masson's trichrome method for histochemical detection of keratin (bright red colour).



primed BALB/c mice were cultured for 9 days in serum-free 'keratinizing medium', containing dibutyl cyclic AMP (10⁻⁴ M), Pap (10⁻⁶ M) as a phosphodiesterase inhibitor¹⁸, and I (5 μ g ml⁻¹). Most (62%) of the mammary glands developed a few foci of epidermoid metaplasia (intensity grade 1) that often differentiated to form keratin (series 3, Table 1). Further, when this medium was supplemented with a mixture of PGE₁, PGE₂ and PGB₁, each at 5 μ g ml⁻¹ (enhancing medium) in series 4, there was considerable enhancement of the ability of dibutyl cyclic AMP to induce the squamous development and differentiation of the epithelium that lined both the mammary ducts and alveoli (Fig. 2a,b). Virtually all glands (99%) contained numerous foci of metaplasia with central cornification and keratin plaques (intensity grade 4). Mitotic figures were present in both the metaplastic and normal epithelial cells. All three PGs were necessary for maximal effect. Lowering the concentration of each PG to 1 μ g ml⁻¹ still invoked the epidermoid metaplasia in all glands, although the foci were fewer and smaller (grade 1–2, not shown). The addition of a mixture of PGF₁ and PGF_{2 α} , each at 5 μ g ml⁻¹, to the keratinizing or enhancing medium did not augment the effect (series 5).

The induction of the epidermoid metaplasia specifically required the presence of cyclic adenine nucleotide. Dibutyl cyclic GMP at the same concentration (10⁻⁴ M), or sodium butyrate (cleavage product of dibutyl cyclic AMP) at twice that concentration, failed to replace the dibutyl cyclic AMP of enhancing medium (11% and 0% frequencies, respectively; series 6 and 7). Cyclic AMP or 5' AMP had little activity compared to the more lipophilic dibutyl cyclic AMP; they required the three PGs to reach the level of activity exhibited by the dibutyl cyclic AMP without PG in keratinizing medium (series 8 and 9). Omission of both dibutyl cyclic AMP and Pap from enhancing medium caused complete loss of activity in the control series 10 (Fig. 2c). A modified enhancing medium lacking only dibutyl cyclic AMP was partially active (30%, intensity grade 1), presumably because of an inhibition of degradation of endogenous cyclic AMP by Pap (series 11). Comparison of results of series 7 and 11 indicates that the presence of sodium butyrate was mildly inhibitory. The high frequency (100%) and intensity (grade 4) of metaplasia following deletion of Pap from enhancing medium are consistent with the presence of only a low level of endogenous cyclic AMP phosphodiesterase activity in the glands (series 12). The use of IPA-H medium or the omission of all supplements except I (regression medium) resulted in no metaplasia (series 1 and 2). Mammary glands of mice not hormonally primed yielded an intermediate level of squamous metaplasia in enhancing medium (each of 26 glands, intensity grade 2).

Table 1 Induction of epidermoid metaplasia in cultured mammary glands

Series	Treatment*	No. of glands	Metaplastic glands	
			Frequency [†] %	Intensity grade [‡]
1	I + P + A + H	30	0	0
2	I	30	0	0
3	Dibutyl cyclic AMP + Pap + I (keratinizing)	31	62	1
4	Dibutyl cyclic AMP + Pap + I + PGE ₁ , E ₂ , B ₁ (enhancing)	114	99	4
5	Dibutyl cyclic AMP + Pap + I + PGE ₁ , E ₂ , B ₁ , F _{1α} , F _{2α}	34	100	4
6	Dibutyl cyclic GMP + Pap + I + PGE ₁ , E ₂ , B ₁	27	11	1
7	Sodium butyrate + Pap + I + PGE ₁ , E ₂ , B ₁	16	0	0
8	Cyclic AMP + I + PGE ₁ , E ₂ , B ₁	21	52	1
9	5'AMP + I + PGE ₁ , E ₂ , B ₁	28	68	1
10	PGE ₁ , E ₂ , B ₁ + I (control)	46	0	0
11	Pap + PGE ₁ , E ₂ , B ₁ + I	27	30	1
12	Dibutyl cyclic AMP + I + PGE ₁ , E ₂ , B ₁	18	100	4

*Whole mammary glands from hormonally primed mice were incubated for 9 d in serum-free medium containing the listed supplements at the following concentrations: dibutyl cyclic AMP, dibutyl cyclic GMP, cyclic AMP, and 5' adenylic acid (5' AMP), each at 10⁻⁴ M; Pap at 10⁻⁶ M; PGs, I, P, A, and H, each at 5 μ g ml⁻¹; and sodium butyrate at 2 $\times$ 10⁻⁴ M.

[†]Per cent of multiple of glands with foci of stratified squamous epithelial cells and keratin deposits.

[‡]Intensity of metaplasia on a scale of 1–4: grade 0, no epidermoid metaplasia; grade 1, minimal epidermoid change. Foci were sparsely distributed and poorly developed. Grades 2 and 3, intermediate grades quantitatively approaching 1 and 4, respectively. Grade 4, well developed foci of epidermoid metaplasia involving the lining of most major ducts and distal ductules. Differentiation of squamous epithelium was partial to complete, the latter forming keratin plaques.

The abnormally developed and differentiated epidermoid cells involution following the removal of the dibutyryl cyclic AMP, Pap and PGs (Fig. 1). Incubation of the metaplastic glands for an additional 15 days (days 9–24) in the regression medium caused complete involution of both the epidermoid and cuboidal (normal secretory) epithelia, leaving behind a residue of keratin plaques (Fig. 2d).

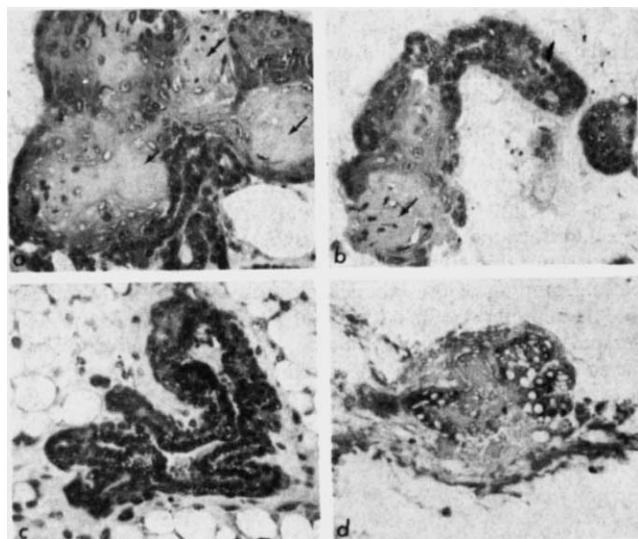


Fig. 2 Microscopic details of squamous metaplastic and involuted mammary glands. Whole glands were fixed, embedded in paraffin, cut at 5 μ m and stained with haematoxylin and eosin⁷. Arrows indicate parakeratotic and keratotic whorls surrounded by layers of differentiating stratified squamous epithelium. *a*, Section of mammary gland that was incubated in enhancing medium for 9 days. Large proximal ducts have extensive epidermoid metaplasia with marked keratosis. Folds of darkly staining cuboidal epithelium (lower centre) are residua of the normal ductal lining pushed into the lumen by the subjacent metaplastic squamous epithelium (series 4, Table 1). *b*, Same preparation showing a distal ductule and alveolar buds with well developed keratinizing epidermoid metaplasia. Normal-appearing cuboidal cells are admixed above and to the right (series 4, Table 1). *c*, Section of a control gland that was incubated in the absence of dibutyryl cyclic AMP and Pap in medium containing I and PGE₁, PGE₂ and PGB₁ for 9 days. The appearance is like that of the primed gland at day 0. The ductal system is developed, but there is no lobuloalveolar growth. The epithelium is of the normal cuboidal type and shows no epidermoid metaplasia (series 10, Table 1). *d*, Section of a gland that was incubated in enhancing medium for 9 days, and then kept in regression medium for 15 days. A distal alveolar lobule that was partially metaplastic has involuted, leaving a residue of vacuolated secretory cells and a central plaque of keratin. All squamous cells have disappeared.

The induction of the metaplasia was not blocked by retinoid. The presence of 2-retinylidene-5,5-dimethyl-1,3-cyclohexanedione (retinylidene dimedone) at 10⁻⁵, 10⁻⁷ or 10⁻⁹ M for 9 days failed to prevent the squamous metaplasia and keratin production brought about by enhancing medium (all 100%, not shown). This retinoid was chosen because it prevents, suppresses and apparently reverses the transformations of the cultured mammary glands caused by procarcinogens^{15,17}. The failure of the retinoid to prevent squamous metaplasia in the present study is in marked contrast to the preventions and reversals of squamous metaplasias arising from vitamin A deficiency^{3,4} and chemical carcinogens⁵⁻⁷ in other systems. If these agents are effective in a common pathway, this difference suggests that the cyclic nucleotide, PGs and Pap act at stages after those of vitamin A deficiency and chemical carcinogens. Alternatively, separate pathways may be involved.

The present findings provide the initial indication of the identities of the possible physiological inducers of epidermoid metaplasia. Cyclic adenine nucleotide, certain PGs and Pap

specifically and reversibly induce this abnormal development and differentiation in the epithelial cells of cultured mammary glands, endowing them with characteristics of skin. The common denominator among the inducers may be that they act synergistically to elevate the level of intracellular cyclic adenine nucleotide in three ways: first, by the exogenous entry of dibutyryl cyclic AMP, second, by the stimulation of the synthesis of cyclic AMP due to the actions of the PGs on adenyl cyclase^{19,20}, and third, through an inhibition by Pap of the phosphodiesterase-mediated degradation of the nucleotide¹⁸.

The normal course of lobuloalveolar development and differentiation (casein production) impeded the abnormal course of squamous metaplastic development and differentiation (keratin production). Prior induction of the normal course in whole mammary glands cultured in medium containing IPAH markedly reduces the subsequent induction of the metaplasia in enhancing medium (F.V.S., R.P.C. and S.S., unpublished). Conversely, agents that elevate the intracellular level of cyclic AMP have recently been reported to inhibit the production of α -lactalbumin and casein by mammary glands in fragment culture²¹.

The possibility arises that cyclic AMP and PGE₁ and PGE₂ may be the mediators of the squamous metaplasias that occur in many abnormal, transformed and malignant tissues, including mammary glands, of animals and humans^{1,7,8,22-26}. Agents that elevate the level of intracellular cyclic AMP have been variously reported either to stimulate or to inhibit the growth of cells of mammary gland²⁷, epidermis²⁸ and other organs^{28,29}. Further, several types of normal differentiation have previously been elicited by these same inducers in cell cultures of fibroblasts³⁰⁻³², skeletal muscle³³, melanoma³⁴ and other systems^{19,20,29}. In addition, the production of keratohyaline granules is stimulated in cultures of human skin by theophylline, an inhibitor of the phosphodiesterase degradation of cyclic AMP³⁵ and keratin filaments are present in the cytoskeletons of epithelial cells in many organs³⁶. It accordingly seems reasonable to investigate whether cyclic adenine nucleotide and specific PGs are the physiological mediators of the normal and abnormal development and differentiation of epithelia into enhanced keratin-producing, stratified squamous cells, as in skin and other organs.

We thank Dr W. Thomas London for critical discussions, Mrs Dovile Cooper for assistance, Mrs Ruth Bender for histological preparations, the Upjohn Company for a gift of PGs and Dr Michael B. Sporn for a gift of retinylidene dimedone, which was synthesized by Drs Nancy Acton and Arnold Brossi. This study was supported in part by NIH grants CA-21522, CA-05945, CA-06927 and RR-05539, and an appropriation from the Commonwealth of Pennsylvania.

Received 14 March; accepted 26 June 1980.

- Anderson, W. A. D. & Kissane, J. M. in *Pathology* Vol. 1, 7th edn, 84–85 (Mosby, St. Louis, 1977).
- McDowell, E. M., Becci, P. J., Schurch, W. & Trump, B. F. *J. natn. Cancer Inst.* **62**, 995–1008 (1979).
- Wolbach, S. B. & Howe, P. R. *J. exp. Med.* **42**, 753–777 (1925).
- Marchok, A. C., Cone, M. V. & Nettesheim, P. *Lab. Invest.* **33**, 451–460 (1975).
- Lasnikski, I. *Br. J. Cancer* **5**, 345–352 (1951).
- Harris, C. C. *et al. J. natn. Cancer Inst.* **48**, 743–761 (1972).
- Tonelli, Q. J., Custer, R. P. & Sorof, S. *Cancer Res.* **39**, 1784–1792 (1979).
- Sporn, M. B., Dunlop, N. M., Newton, D. L. & Smith, J. M. *Fedn Proc.* **35**, 1332–1338 (1976).
- Ichinose, R. R. & Nandi, S. *J. Endocr.* **35**, 331–340 (1966).
- Wood, B. G., Washburn, L. L., Mukherjee, A. S. & Banerjee, M. R. *J. Endocr.* **65**, 1–6 (1975).
- Terry, P. M., Bail, E. M., Ganguly, R. & Banerjee, M. R. *J. immun. Meth.* **9**, 123–134 (1975).
- Tonelli, Q. J. & Sorof, S. *Nature* **285**, 250–252 (1980).
- Banerjee, M. R., Wood, B. G. & Washburn, L. L. *J. natn. Cancer Inst.* **53**, 1387–1393 (1974).
- Telang, N. T., Banerjee, M. R., Iyer, A. P. & Kundu, A. B. *Proc. natn. Acad. Sci. U.S.A.* **76**, 5886–5890 (1979).
- Dickens, M. S., Custer, R. P. & Sorof, S. *Proc. natn. Acad. Sci. U.S.A.* **76**, 5891–5895 (1979).
- Kundu, A. B., Telang, N. T. & Banerjee, M. R. *J. natn. Cancer Inst.* **61**, 465–469 (1978).
- Dickens, M. S. & Sorof, S. *Nature* **285**, 581–584 (1980).
- Prasad, K. N. & Kumar, S. in *Control of Proliferation in Animal Cells* (eds Clarkson, B. & Baserga, R.) 581 (Cold Spring Harbor Laboratory, New York, 1974).
- Samuelsson, B., Granström, E., Green, K., Hamberg, M. & Hammarström, S. A. *Rev. Biochem.* **44**, 669–695 (1975).
- Samuelsson, B. *et al. A. Rev. Biochem.* **47**, 997–1029 (1978).
- Perry, J. W. & Oka, T. *Proc. natn. Acad. Sci. U.S.A.* **77**, 2093–2097 (1980).

22. Pasternack, J. G. & Wirth, J. E. *Am. J. Path.* **12**, 423-437 (1936).
23. Deringer, M. K. *J. natn. Cancer Inst.* **22**, 995-1002 (1959).
24. Huvos, A. G., Lucas, J. C. & Foote, F. W. *N. Y. St. J. Med.* **73**, 1078-1082 (1973).
25. Medina, D. *J. natn. Cancer Inst.* **53**, 213-221 (1974).
26. Hassan, M. O. & Olaiyola, M. Y. *Archs Path. Lab. Med.* **103**, 624-630 (1979).
27. Taylor-Papadimitriou, J., Purkis, P. & Fentiman, I. S. *J. cell. Physiol.* **102**, 317-321 (1980).
28. Green, H. *Cell* **15**, 801-811 (1978).
29. Friedman, D. L. *Physiol. Rev.* **56**, 652-708 (1976).
30. Johnson, G. S., Friedman, R. M. & Pastan, I. *Proc. natn. Acad. Sci. U.S.A.* **68**, 425-429 (1971).
31. Hsie, A. W. & Puck, T. T. *Proc. natn. Acad. Sci. U.S.A.* **68**, 358-361 (1971).
32. Sheppard, J. R. *Proc. natn. Acad. Sci. U.S.A.* **68**, 1316-1320 (1971).
33. Aw, E. J., Holt, P. G. & Simons, P. J. *Expl Cell Res.* **83**, 436-438 (1973).
34. Johnson, G. S. & Pastan, I. *Nature new Biol.* **237**, 269-270 (1972).
35. Chopra, D. *Br. J. Derm.* **96**, 255-262 (1977).
36. Sun, T.-T., Shih, C. & Green, H. *Proc. natn. Acad. Sci. U.S.A.* **76**, 2813-2817 (1979).

Specific binding of ^3H -substance P to rat brain membranes

M. R. Hanley*, B. E. B. Sandberg*, C. M. Lee*, L. L. Iversen*, D. E. Brundish† & R. Wade†

*MRC Neurochemical Pharmacology Unit, Department of Pharmacology, University of Cambridge Medical School, Hills Road, Cambridge CB2 2QD, UK

†Ciba-Geigy Pharmaceuticals Division, Horsham, West Sussex RH12 4AB, UK

The undecapeptide substance P is a putative neurotransmitter in the mammalian central nervous system (CNS), and may be associated with pain fibres in the spinal cord^{1,2}. Radiolabelled derivatives of other neuropeptides have been used to demonstrate specific interactions with receptor sites on brain membranes³⁻⁶, and this approach has now been explored with substance P. We have now prepared [^3H -Phe⁸]-substance P and we find that it binds reversibly to a saturable population of sites in rat brain particulate fractions. Scatchard analysis of concentration-dependent saturation of binding indicates a single population of non-interacting sites with a high affinity ($K_d = 0.38 \text{ nM}$) and a low density ($B_{\text{max}} = 27.2 \text{ fmol per mg protein}$). Kinetic analyses indicate an apparent dissociation equilibrium constant of 0.46 nM . A variety of neurotransmitter amines and amino acids, and other peptides do not compete at the substance P sites, but structurally related peptides or shorter C-terminal fragments of substance P are active. The rank order of potency of these substance P-related peptides agrees with that reported for their effects in depolarizing spinal cord neurones⁷. The regional distribution of the specific binding sites for ^3H -substance P parallels that of substance P immunoreactivity, being high in the hypothalamus and low in the cerebellum and cerebral cortex. The characteristics of the ^3H -substance P binding sites are consistent with those expected for substance P receptors.

^3H -labelled substance P was prepared by reductive deiodination of the protected peptide [Boc-Arg¹,Lys(Boc)³,Phe(I)⁸]-substance P with tritium as similarly described for corticotropin-(1-24)-tetracosapeptide⁸. The solid-phase synthesis of the precursor peptide and the tritiation and purification of the radiolabelled substance P will be reported in detail elsewhere. The specific activity of the final product was 23 Ci mmol^{-1} , based on quantitative amino acid analysis and confirmed by radioimmunoassay. The purity was assessed by TLC in three different solvent systems, by isocratic and gradient-elution reverse-phase HPLC in three systems, and by measurements of the immunoreactivity of serially diluted samples against C-terminal sequence- and N-terminal sequence-directed⁹ antisera. By all techniques, the purity was $\geq 90\%$, and the major contaminant was substance P sulfoxide, in which the C-terminal methionine residue was oxidized¹⁰. Aliquots of ^3H -substance P were stored at a concentration of 0.25 mM in 0.01 M HCl in glass vials under liquid nitrogen. For use, samples were diluted to $25 \mu\text{M}$ with 0.1 M HCl containing 1 mg ml^{-1} bovine serum albumin (BSA; recrystallized, Sigma).

In these conditions, the labelled peptide could be stored at -20°C for up to 10 days, with repeated freezing and thawing, and accumulated $<5\%$ of ^3H -substance P sulfoxide and negligible amounts of other labelled products. On the guinea pig ileum, ^3H -substance P had contractile activity equipotent with that of synthetic unlabelled substance P.

^3H -substance P was rapidly degraded when incubated at 37°C with a crude synaptic membrane preparation from rat brain (see Fig. 1 legend for procedure). For this reason, all incubations included the peptidase inhibitors bacitracin ($300 \mu\text{g ml}^{-1}$), phenylmethylsulphonyl fluoride (0.1 mM), and *p*-chloromercuribenzoate (0.1 mM), and the incubations were performed at 4°C . ^3H -substance P was stable to enzymatic digestion in these conditions as determined by HPLC analyses of acid extracts of membrane-bound radioactivity after various intervals (5-60 min). Control experiments indicated that none of these additions interfered with ^3H -substance P binding.

At low concentrations of ^3H -substance P ($<1 \text{ nM}$), 50-60% of the total binding of the labelled peptide to brain membranes could be inhibited by the addition of $10 \mu\text{M}$ unlabelled peptide. The specific component of binding, defined in this way, was saturable, whereas the nonspecific binding (in the presence of $10 \mu\text{M}$ substance P) was linearly dependent on ligand concentration (Fig. 1a). The specific and nonspecific components were both linearly dependent on the amount of tissue present in the

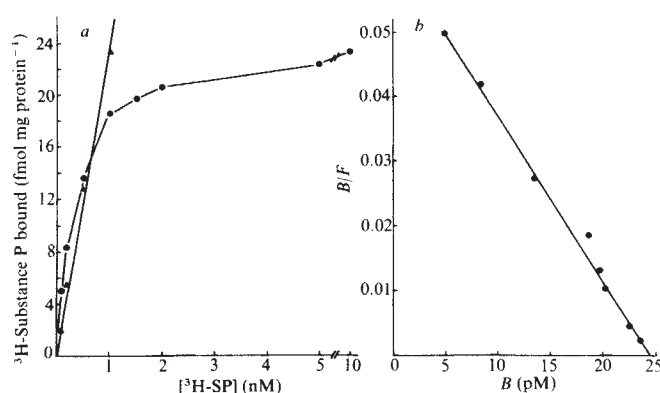


Fig. 1 a, Concentration dependence of specific (circles) and nonspecific (triangles) ^3H -substance P binding to rat brain membrane fractions. Whole brains minus cerebral cortex and cerebellum were dissected on ice from adult male Sprague-Dawley rats and homogenized in 10 vol ice-cold 10% sucrose/20 mM HEPES (pH 7.4) with a Teflon-glass homogenizer. The homogenate was centrifuged at $1,000g$ (5 min, 4°C) to remove cellular debris. The supernatant was centrifuged at $11,000g$ (20 min, 4°C) to prepare a crude mitochondrial pellet. The pellet was lysed by resuspension in 10 vol of ice-cold 5 mM HEPES (pH 8.1). The lysate was centrifuged at $100,000g$ (30 min, 4°C). The resulting pellet was resuspended with a Teflon-glass homogenizer in 5 vol of 20 mM HEPES (pH 7.4) and stored frozen at -20°C . For binding assays, $50 \mu\text{l}$ of the thawed membrane suspension were added to $430 \mu\text{l}$ of incubation medium (20 mM HEPES pH 7.4 containing 0.5% recrystallized BSA, $300 \mu\text{g ml}^{-1}$ bacitracin, 0.1 mM phenylmethylsulphonyl fluoride, 0.1 mM *p*-chloromercuribenzoate) in small plastic tubes kept on ice. The tubes were preincubated for 60 min before further additions. Drug, radiolabel and blank additions were made in $10\text{-}\mu\text{l}$ aliquots so that the final volume was $500 \mu\text{l}$. Nonspecific binding was defined by adding substance P to a final concentration of $10 \mu\text{M}$. In all cases, total and nonspecific binding determinations were made in triplicate. The ^3H -substance P was diluted from a $25 \mu\text{M}$ stock solution with incubation medium. Incubations were performed for 20 min on ice, and were terminated by dilution with 2.5 ml of ice-cold 20 mM HEPES/1% BSA (HEPES/BSA). Membrane-bound ^3H -substance P was measured by rapid filtration of the diluted samples through Millipore EGWP ($0.2 \mu\text{m}$) filters pre-equilibrated with ice-cold HEPES/BSA for 3 h. Filters were rinsed with $3 \times 2.5 \text{ ml}$ of ice-cold HEPES/BSA. Filtration and rinsing were completed in $<1 \text{ min}$. Filters were clarified by the addition of 4 ml ethoxyethanol and counted by the addition of 10 ml toluene scintillant. The data points are the means of triplicates in a single experiment. These results have been replicated in four separate experiments by the same technique. In three additional experiments using a centrifugation assay to measure binding, the specific binding data were reproduced, but nonspecific binding represented a greater proportion of the total (70-80%). b, Scatchard plot of the specific binding data in a.

incubation. A Scatchard analysis (Fig. 1b) of the saturation curve for specific binding yielded a single straight line from which the density of the sites ($B_{\max}$) was 27.2 fmol per mg protein and the dissociation constant (K_d) was calculated as 0.38 nM. Addition of a large molar excess of unlabelled substance P to membranes previously incubated with ^3H -substance P led to an exponential decay of specific binding (Fig. 2b). These data, when analysed by a semi-log plot, gave a straight line with a slope of -0.15 min^{-1} , corresponding to a half time of dissociation of 4.6 min, and a dissociation rate constant (k_{off}) of $2.5 \times 10^{-3} \text{ M}^{-1} \text{ s}^{-1}$ (Fig. 2b). The time course for association of 2 nM ^3H -substance P reached a plateau for nonspecific sites at 2 min (data not shown) and for specific sites at 10 min (Fig. 2a). Plotting these data according to a first-order reaction gave a straight line with a slope of 0.47. The association rate constant (k_{on}) determined from the slope was $5.4 \times 10^6 \text{ M}^{-1} \text{ s}^{-1}$.

Table 1 Inhibition of specific ^3H -substance P binding to rat brain membranes by related peptides

Peptide	Rat brain binding inhibition $\text{IC}_{50}(\text{nM})^*$	Relative potency	Rat spinal cord depolarization (ref. 7) (relative potency)
Substance P (SP)	3	1.0	1.0
SP ₂₋₁₁	10	0.3	0.6–0.9
SP ₃₋₁₁	36	0.08	0.4–1.0
SP ₄₋₁₁	64	0.05	0.8–1.0
SP ₅₋₁₁	6	0.5	2–12
SP ₆₋₁₁ [†]	1.8	1.7	5–12
SP ₇₋₁₁	220	0.01	<0.02
SP free acid	100	0.03	ND
Eledoisin	11	0.3	ND
Physalaemin	1.6	1.9	ND

* Determined from Hill plots of inhibition of binding of 2 nM ^3H -substance P. Data were generated from the means of triplicates of at least three separate experiments. For experiments with C-terminal fragments of substance P, 1 mM puromycin was added to the incubation medium to protect against aminopeptidase degradation. All peptides were commercial products (Peninsula) and were checked for purity by HPLC.

[†] Pyroglutamyl N-terminal form. Negative at 10 μM : angiotensin II and III, apamin, bombesin, bradykinin, α -bungarotoxin, carnosine, cholecystokinin-8 (non-sulphated), cholecystokinin-33, leucine-enkephalin, methionine-enkephalin, β -endorphin, β -lipotropin, neurotensin, SQ-20881, thyrotropin-releasing hormone, vasoactive intestinal peptide, vasopressin. Negative at 100 μM : aspartic acid, baclofen, capsaicin, carbamylcholine, dopamine, γ -aminobutyric acid, glutamic acid, glycine, histamine, serotonin, noradrenaline.

Using the equation $K_d = k_{\text{off}}/k_{\text{on}}$, an equilibrium dissociation constant of 0.46 nM was calculated, in good agreement with that determined from saturation data.

The sites were inhibited by structurally related analogues and by C-terminal fragments of substance P but not by a range of other neurotransmitters or peptides (Table 1). The rank order of potencies for a series of C-terminal fragments of substance P was hexapeptide > substance P > heptapeptide > decapeptide > nonapeptide > octapeptide > pentapeptide. Physalaemin and eledoisin, naturally-occurring substance P-related peptides¹¹, were also effective in competing for specific sites. Substance P free acid, in which the C-terminal methionine amide is replaced by a methionine, was about 30 times less active than substance P. Baclofen, a putative substance P antagonist⁷, had no effect on ^3H -substance P binding. The IC_{50} for nonradioactive substance P (against 2 nM ^3H -substance P) was 3 nM, corresponding to an apparent K_d of 0.48 nM, in close agreement with both kinetic and equilibrium measurements of the dissociation constant.

A preliminary survey of the regional distribution of ^3H -substance P binding in rat brain was performed, using a single radioligand concentration (2 nM). ^3H -substance P binding sites were unevenly distributed among rat brain regions (Table 2). The density of specific sites was highest in substance P-rich areas

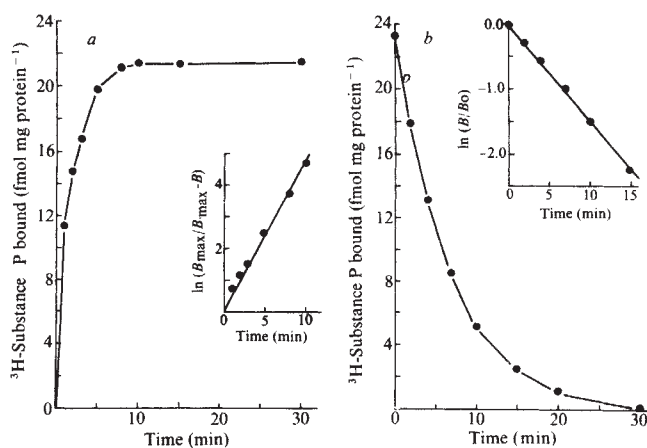


Fig. 2 a, Time course of specific ^3H -substance P binding at 4 °C. 2 nM ^3H -substance P was incubated with rat brain membrane fractions, and total and nonspecific binding were determined at the indicated intervals to generate the specific binding. The points are the means of triplicates in three separate experiments. Insert: Linearization of the time course curve for calculation of k_{on} as described elsewhere⁶. b, Dissociation time course of specific ^3H -substance P binding at 4 °C. 2 nM ^3H -substance P was incubated on ice for 20 min with rat membrane fractions. Substance P was then added to a final concentration of 10 μM at time zero. Samples of total and nonspecific binding were filtered at the indicated intervals afterwards. The results are the means of triplicates in three separate experiments. Insert: Linearization of the dissociation curve for calculation of k_{off} as described elsewhere⁶.

such as hypothalamus and midbrain, and lowest in cerebral cortex and cerebellum, which contain only small amounts of substance P¹².

The structural specificity of the high affinity binding sites in rat brain resembles that determined by measurements of the biological activity of substance P and related peptides¹³. Most studies have focused on the actions of substance P on peripheral tissues, but Otsuka and Konishi⁷ examined the potencies of substance P and shorter fragments in depolarizing motor neurones after bath application to newborn rat spinal cord, and their results (summarized in Table 1) are in agreement with the present binding data. In particular, it is noteworthy that the hexapeptide fragment is more potent than the native peptide, both in its activity on spinal cord and in inhibiting ^3H -substance P binding, whereas the C-terminal pentapeptide fragment shows a marked reduction in activity. Pierce and Einspahr¹⁴ have also shown a marked reduction in activity of the C-terminal pentapeptide electrophysiologically on dorsal horn neurones.

Table 2 Regional distribution of specific ^3H -substance P binding to rat brain membrane fractions

Region	Binding (fmol per mg protein)
Brainstem	15 ± 5
Cerebellum	6 ± 3
Hypothalamus	36 ± 9
Midbrain	29 ± 6
Striatum	17 ± 4
Temporal cortex	4 ± 2

Individual regions were dissected on ice and crude mitochondrial fraction membranes prepared as in Fig. 1. The specific binding of 2 nM ^3H -substance P was determined using membrane fractions stored frozen at -20°C as described in Fig. 1 legend. The protein content was determined by the Lowry technique. The results are means $\pm$ s.e.m. for triplicate measurements of total and nonspecific binding in four separate experiments. In preliminary experiments substance P-displaceable binding was observed in the cerebellum, cerebral cortex and hippocampus that could also be inhibited by angiotensin II and SQ-20881 (0.1–10 μM), unlike the binding in other regions or in total brain (minus cerebellum and cortex). The nature of this unselective peptide-binding site is not known, but the incubation medium was routinely supplemented with 10 μM angiotensin II to eliminate its contribution.

Although data for the substance P-related peptides physalaemin and eledoisin are not available for rat spinal cord, these substances are approximately equipotent with substance P in a number of other test systems^{13,15}, and they had a similar potency to substance P in the binding assay. Removal of the C-terminal amide in substance P free acid greatly reduces biological potency in a number of test systems^{13,16}, and was also found markedly to reduce potency in the binding assay.

Although there have been other published reports of high-affinity binding of radioactively labelled substance P derivatives to brain membrane preparations^{17,18}, we believe that the present results reflect most closely the known biological activity of substance P. Thus, Nakata *et al.*¹⁷ described the binding of ³H-substance P to rabbit brain membrane preparations, but observed no antagonism by physalaemin. Furthermore, we have been unable to replicate their measurement technique as we observe large amounts of ³H-substance P binding to the type of membrane filters used by these authors, even in the absence of tissue. From the present results, it seems unlikely that the specific binding of ³H-substance P could have reached equilibrium during the 1 min incubation period at 0 °C used by these authors. Mayer *et al.*¹⁸ studied the binding of [¹²⁵I-Tyr⁸]-substance P to a synaptic vesicle fraction from rat brain. Although the labelled peptide exhibited high affinity binding ($K_d = 0.32$ nM), and the regional distribution of binding sites in brain resembled that of endogenous substance P¹⁹, the specificity of this binding bore little relation to biological activity, since C-terminal fragments of substance P inhibited binding only at micromolar concentrations, and physalaemin enhanced rather than inhibited binding of the labelled analogue. Putney *et al.*²⁰ and Jensen and Gardner²¹ have described the use of [¹²⁵I]-physalaemin to label substance P receptors in dispersed parotid acinar cells and dispersed pancreatic acinar cells respectively. In each case, the authors described specific high-affinity binding sites that were present in unusually low density. It is interesting that the density of ³H-substance P sites in rat brain is lower than that seen for many conventional amine or amino acid neurotransmitter receptor sites²², but is similar to the values reported for other brain peptide-binding sites³⁻⁵.

The characteristics and regional distribution of the high affinity binding sites for ³H-substance P described here are in close correspondence with those expected for substance P receptors in the CNS. The radioligand binding assay offers an important new tool for investigating the occurrence and pharmacological properties of such substance P receptors.

M.R.H. is a Fellow of the Helen Hay Whitney Foundation.
C.M.L. is a Li Po Chun Scholar.

Received 28 April; accepted 23 June 1980.

- Cuello, A. C. *et al.* in *Centrally Acting Peptides* (ed. Hughes, J.) 135-156 (Macmillan, London, 1978).
- Piercey, M. F., Einspahr, F. J., Dobby, P. J. K., Schroeder, L. A. & Hollister, R. P. *Brain Res.* **186**, 421-434 (1980).
- Moody, T. W., Pert, C. B., Rivier, J. & Brown, M. R. *Proc. natn. Acad. Sci. U.S.A.* **75**, 5372-5376 (1978).
- Sirett, N. E., McLean, A. J., Bray, J. J. & Hubbard, J. I. *Brain Res.* **122**, 299-312 (1977).
- Taylor, D. P. & Pert, C. B. *Proc. natn. Acad. Sci. U.S.A.* **76**, 660-664 (1979).
- Kitagawa, K. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **74**, 1846-1850 (1977).
- Otsuka, M. & Konishi, S. *Cold Spring Harb. Symp. quant. Biol.* **40**, 135-143 (1976).
- Brundish, D. E. & Wade, R. J. *Chem. Soc. Perkin Trans. I*, 2186-2189 (1973).
- Lee, C. M., Emson, P. C. & Iversen, L. L. *Life Sci.* (in the press).
- Floor, E. & Leeman, S. *Analyt. Biochem.* **101**, 498-503 (1980).
- Bertaccini, G. *Pharmac. Rev.* **28**, 127-177 (1976).
- Kanazawa, I. & Jessell, T. *Brain Res.* **117**, 362-367 (1976).
- Bury, R. W. & Mashford, M. L. *Aust. J. exp. Biol. med. Sci.* **55**, 671-735 (1977).
- Piercey, M. F. & Einspahr, F. J. *Brain Res.* **187**, 481-486 (1980).
- Ersapamer, G. F., Ersapamer, V. & Piccinelli, D. *Naunyn-Schmiedeberg Archs Pharmac.* **311**, 61-65 (1980).
- Couture, R., Fournier, A., Magnan, J., St. Pierre, S. & Regoli, D. *Can. J. Physiol. Pharmac.* **57**, 1427-1436 (1979).
- Nakata, Y., Kusaka, Y., Segawa, T., Yajima, H. & Kitagawa, K. *Life Sci.* **22**, 259-268 (1978).

- Mayer, N., Lembeck, F., Saria, A. & Gamse, R. *Naunyn-Schmiedeberg Archs Pharmac.* **306**, 45-51 (1979).
- Saria, A., Mayer, N., Lembeck, F. & Pabst, M. *Naunyn-Schmiedeberg Archs Pharmac.* **311**, 151-157 (1980).
- Putney, J., Van De Walle, C. & Wheeler, C. J. *Physiol., Lond.* **301**, 205-212 (1980).
- Jensen, R. T. & Gardner, J. D. *Proc. natn. Acad. Sci. U.S.A.* **76**, 5679-5683 (1979).
- Snyder, S. H. & Bennett, J. P. A. *Rev. Physiol.* **38**, 153-175 (1976).

β -Endorphin alters luteinizing hormone secretion via the amygdala but not the hypothalamus

N. Parvizi & F. Ellendorff

Institut für Tierzucht und Tierverhalten, FAL, Mariensee, 3057 Neustadt 1, FRG

Morphine and enkephalins are able to alter pituitary hormone secretion¹⁻³. It has been postulated that they do not act directly at the pituitary⁴, but rather that the hypothalamus is the site at which inhibition or stimulation of pituitary hormone secretion is initiated⁴. On the other hand, endogenous opiates have been located in distinctly different neuronal regions, including areas outside the hypothalamus⁵. The effects of β -endorphin on plasma luteinizing hormone (LH) levels have not been explored, and the present experiments attempt to elucidate the contribution and the possible site of action of β -endorphin in the control of LH secretion. The results show that β -endorphin inhibits pituitary LH secretion if applied into the amygdala but not when given into the hypothalamus.

Stainless steel guide tubings (22 gauge) containing inner 28-gauge tubing were implanted bilaterally in the basolateral

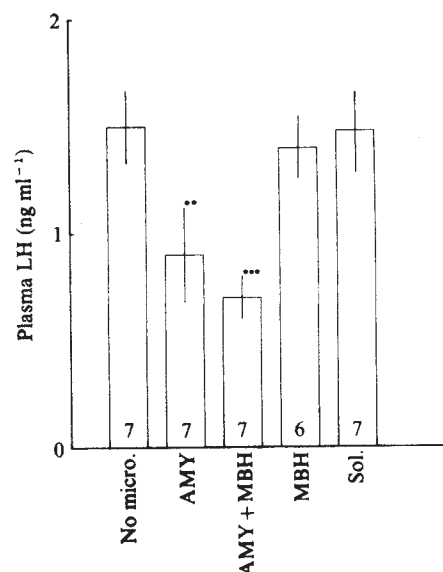


Fig. 1 Decrease of plasma LH levels 100-180 min after microinjection of 30 ng β -endorphin into the amygdala (AMY) or medio-basal hypothalamus (MBH) and amygdala of ovariectomized miniature pigs. 'No micro.', control blood sampling without any treatment; 'Sol.', microinjection of 1 μ l of 0.7% saline. Numbers in the bars represent the number of animals. For statistical analysis the period of sampling was divided into three blocks each of nine successive samples taken at 10-min intervals. The mean LH value for each animal within each block was calculated. A one-way analysis of variance was then carried out to detect differences among mean LH values of treatments and time blocks. ** $P \leq 0.01$; *** $P \leq 0.001$; $\bar{x} \pm$ s.e.m.

hypothalamus and amygdala of seven adult, ovariectomized miniature pigs. The tips of the tubing ended 3 mm above the desired depth. Two weeks later and 5 days before the beginning of the experiment, a silastic catheter was placed in the external jugular vein. During the recovery period, the animals were familiarized with the manipulations of the implant. The experiment was designed so that each animal served as its own control. A control period preceded the treatment on each

experimental day. Saline (solvent), 1 μ l, β -endorphin (Serva cat. no. 52558), 30 ng, were microinjected on 4 different days at 3-day intervals into either the basolateral hypothalamus or amygdala, or simultaneously into the two tissues. The sequence of treatment was randomly assigned for each animal. Microinjections were given through a 28-gauge tubing connected to a Hamilton syringe as described^{6,7}. Blood samples (2 ml) were collected at 10-min intervals from 90 min before to 180 min after microinjection. On one day, blood samples were taken without any treatment. Thus, a double control system allowed us to consider the animals as a group and individually. Plasma LH was estimated by radioimmunoassay⁸.

Histological verification at the end of the experiment indicated that in six out of seven animals the tips of the tubing were located in the basolateral hypothalamus, and in all seven animals the tips were within the amygdala. Microinjection of β -endorphin into the basolateral hypothalamus had no significant effect on LH secretion, but microinjection into the amygdala of the same animals decreased plasma LH levels significantly ($P \leq 0.01$; for statistical analysis see Fig. 1 legend). The onset of inhibition occurred within 60 min of microinjection (Fig. 2). The simultaneous microinjection of β -endorphin into the hypothalamus and amygdala resulted in a decrease in plasma LH levels similar to that seen with microinjection into the amygdala alone (Fig. 2). This would be expected if β -endorphin were priming the hypothalamus to incoming signals.

The application of other substances, such as steroids^{6,7,9} or catecholamines⁹, directly into the basolateral hypothalamus causes dramatic changes in LH secretion. Hence the absence of any changes in plasma LH after β -endorphin microinjection into the hypothalamus was unexpected. It suggests, among other possibilities, that hypothalamic LH-releasing hormone (LH-RH) secreting cells are not directly sensitive to β -endorphin and that β -endorphin-induced changes in LH originate from extra-hypothalamic neurones, thereby in some way inhibiting LH-RH. Thus, the high number of β -endorphin receptors in the hypothalamus¹⁰ must be involved in mechanisms not related to LH-RH secretion. The amygdala, on the other hand, must now be considered as one site where β -endorphins, among other substances, induce changes that eventually result in altered LH-RH discharge and thus inhibition of pituitary LH secretion. Although we do not know how this is brought about, the amygdala has been identified as a preferred site of endogenous opiate accumulation^{5,10}. However, the effect on plasma LH that we have observed is probably only one effect of amygdala- β -endorphin interplay. For example, it is known that the amygdala and β -endorphin^{12,13}, as well as LH-RH^{14,15}, are involved in controlling a wide range of behavioural patterns.

This work was supported by the Deutsche Forschungsgemeinschaft. We thank Mrs A. Schulte-Derne and Miss R. Duensing for assistance.

Received 24 January; accepted 18 June 1980.

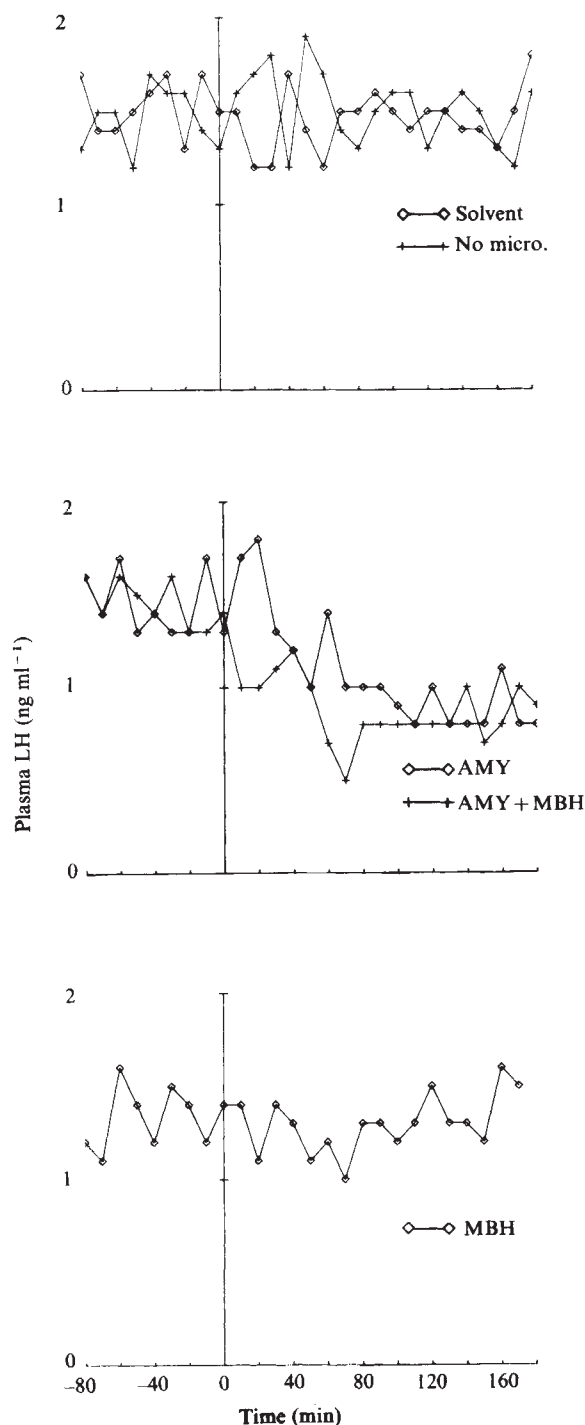


Fig. 2 Point by point presentation of plasma LH levels before and after microinjection of 30 ng β -endorphin into the mediobasal hypothalamus (MBH), amygdala (AMY) or both tissues. 'Solvent' indicates microinjection of 1 μ l 0.7% saline into the hypothalamus and amygdala at time 0; 'No micro' indicates control blood sampling without any microinjection; each point represents seven (AMY, control or solvent) or six (MBH, AMY + MBH) animals. For clarity the s.e.m. values are omitted.

- Dupont, A., Cusan, L., Garon, M., Labrie, F. & Li, C. H. *Proc. natn. Acad. Sci. U.S.A.* **74**, 358-359 (1977).
- Jeiri, T., Chen, H. T. & Meites, J. *Neuroendocrinology* **29**, 288-292 (1979).
- Bruni, J. F., van Vugt, D., Marshall, S. & Meites, J. *Life Sci.* **21**, 471-466 (1977).
- Meites, J., Bruni, J. F., van Vugt, D. A. & Smith, A. F. *Life Sci.* **24**, 1325-1336 (1979).
- Bloom, F., Ballenberg, E., Rossier, J., Ling, N. & Guillemin, R. *Proc. natn. Acad. Sci. U.S.A.* **75**, 1591-1595 (1978).
- Parvizi, N. & Ellendorff, F. *Nature* **256**, 59-60 (1975).
- Parvizi, N., Elsaesser, F., Smidt, D. & Ellendorff, F. *Endocrinology* **101**, 1078-1087 (1977).
- Pomerantz, D. K., Ellendorff, F., Elsaesser, F., König, A. & Smidt, D. *Endocrinology* **94**, 330-335 (1974).
- Parvizi, N. & Ellendorff, F. *Brain Res.* **148**, 521-525 (1978); *J. Steroid Biochem.* **12**, 331-335 (1980).
- Uhl, G. R., Childers, S. R. & Synder, S. H. in *Frontiers in Neuroendocrinology* Vol. 5 (eds Ganong, W. F. & Martini L.) 289-328 (Raven, New York, 1978).
- Eleftherios, B. G. (ed.) *The Neurobiology of the Amygdala* (Plenum, New York, 1972).
- Bloom, F., Segal, D. & Guillemin, R. *Science* **194**, 630-632 (1976).
- Tache, Y., Lis, M. & Collu, R. *Life Sci.* **21**, 841-846 (1977).
- Moss, R. L. & McCann, S. M. *Science* **181**, 177-179 (1973).
- Pfaff, D. W. *Science* **182**, 1148-1149 (1973).

Genetic alteration of nerve membrane excitability in temperature-sensitive paralytic mutants of *Drosophila melanogaster*

Chun-Fang Wu*†‡ & Barry Ganetzky*§

* Division of Biology, California Institute of Technology, Pasadena, California 91125

† Department of Zoology, University of Iowa, Iowa City, Iowa 52242

§ Laboratory of Genetics, University of Wisconsin, Madison, Wisconsin 53706

Single mutations that disrupt basic physiological processes such as axonal conduction and synaptic transmission provide one means to study the membrane components and mechanisms that underlie these processes¹⁻⁸. Temperature-sensitive paralytic mutants of *Drosophila melanogaster* which behave normally at a permissive temperature (23–25 °C) but become paralysed at a restrictive temperature (usually 29–37 °C) have been utilized in such studies. An example is *nap^{ts}* (no action potential, temperature-sensitive), a recessive second-chromosome mutation that causes paralysis of larvae and adults at 37 °C. Electrophysiological recordings have demonstrated that *nap^{ts}* specifically affects a component of axonal membranes⁵. Here we show that axonal conduction in *nap^{ts}* larvae is abnormally sensitive to tetrodotoxin (TTX), a specific blocker of voltage-sensitive sodium channels⁹, and that the refractory period of action potentials differs from wild type, consistent with the idea that *nap^{ts}* alters the function of sodium channels. It is further shown that *para^{ts}* (paralysed, temperature-sensitive), a recessive X-linked mutation¹, also specifically disrupts axonal conduction and that the two mutations affect functionally related membrane components. Identification of other mutants defective in associated components of nerve membranes is made feasible by using techniques described here.

The TTX sensitivity of nerve conduction was measured in both mutant and normal (Canton-S) larvae. The postsynaptic excitatory junctional potentials (e.j.ps) were recorded intracellularly from medioventral longitudinal muscles in the larval neuromuscular preparation^{5,10}. In *Drosophila*¹⁰ and other dipteran^{13,14} larvae, these muscles are innervated by two multi-terminal axons with multiple branches and lack regenerative action potentials under physiological conditions¹⁰. Blockage of axonal conduction causes the loss of the evoked e.j.p. For the *nap^{ts}* larva shown in Fig. 1a, 5 nM TTX has diminished the e.j.p. appreciably and at 10 nM the e.j.p. is completely gone. Axonal conduction but not synaptic transmission is blocked because a graded postsynaptic potential with short latency distinct from the normal e.j.p. can be evoked electrotonically (Fig. 1b). The wild-type e.j.p. is only slightly affected by 10 nM TTX, diminishing at 20 nM TTX, and finally disappearing at 40 nM TTX, at which point a graded response can be evoked electrotonically (Fig. 1b).

In any given larva, we found little variation in the dose of TTX required to block the e.j.p. in different body segments. For the different *nap^{ts}* larvae tested, the effective concentration ranged from 5–15 nM (9.8 ± 4.4 nM; $\bar{x} \pm S.D.$; $n = 11$ larvae). For wild type the range was 20–65 nM (40.3 ± 11.1 nM; $n = 9$ larvae). The difference in TTX sensitivity was confirmed by direct recording of nerve compound action potential. This result also indicates that *nap^{ts}* affects the TTX sensitivity of both the sensory and motor axons in the nerve bundle.

The observed difference in TTX sensitivity is not a consequence of differences in genetic background other than the *nap^{ts}* mutation responsible for the paralytic phenotype. By using a translocation which inserts the *nap⁺* locus into the Y

chromosome⁵, the TTX sensitivity of *X/Y; nap^{ts}/nap^{ts}* and *X/Ynap⁺; nap^{ts}/nap^{ts}* larvae were compared. These larvae are essentially identical in genetic background, but the former have the behavioural phenotype of *nap^{ts}* whereas the latter are *nap⁺*. The e.j.p. in these control larvae was blocked at a TTX concentration of 32.5 ± 10.9 nM ($n = 4$ larvae), comparable to the effective dose for Canton-S (40.3 ± 11.1 nM).

The altered TTX sensitivity is consistent with the idea that *nap^{ts}* alters the function or number of sodium channels⁵. Biochemical analysis of nerve membrane properties in *Drosophila* has begun^{7,15} and can provide independent evidence. Preliminary results from competitive, saturable binding of ³H-TTX to a crude particulate fraction prepared from adult heads indicate different binding properties between normal and *nap^{ts}* extracts¹⁵.

A possible alteration in the density and/or kinetics of voltage-dependent ionic channels in *nap^{ts}* is also suggested by an abnormal refractory period of the compound action potential. The axonal refractory period is determined by the kinetics of activation of sodium and potassium channels and the inactivation of sodium channels¹⁶. For each axon in a nerve bundle there is a definite refractory period during which a second stimulus cannot elicit another propagated response. The second compound action potential gradually recovers as the interval between two stimuli increases, because an increasing number of axons are again able to generate a propagated response (Fig. 2a). The second response recovers more rapidly in control than in *nap^{ts}* larvae even though the stimulus strength relative to the threshold is the same. We defined $\tau_{1/2}$ as the interval between stimuli needed to obtain a second response with amplitude 50% of the first. The $\tau_{1/2}$ decreases for both mutant and control larvae as the stimulus intensity increases (Fig. 2b). However, at almost every relative stimulus strength, the mutant requires a longer interval for recovery than does the control. The minimum $\tau_{1/2}$ is

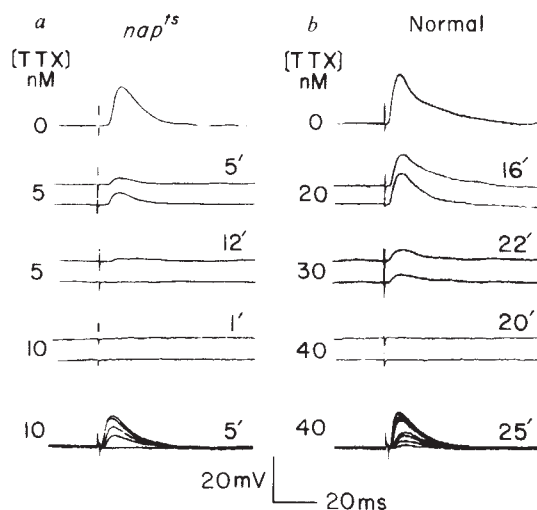
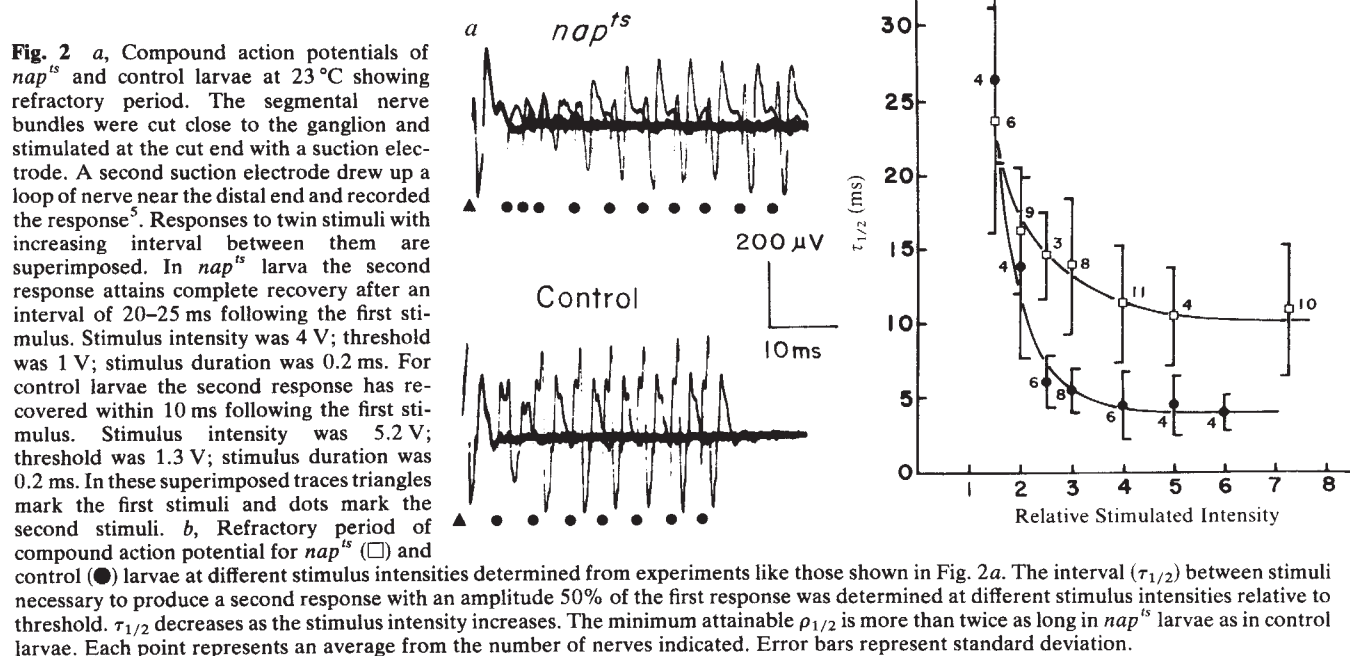


Fig. 1 Continuous recordings (23 °C) from ventrolateral longitudinal muscles in the mutant *nap^{ts}* and the normal (Canton-S) larvae showing e.j.ps at different TTX concentrations shown to the left of each trace. Segmental nerve bundles were severed close to the ventral ganglion and stimulated (0.2 ms duration) via a glass suction electrode. E.j.ps were recorded intracellularly using glass microelectrodes filled with 3M KCl (30–60 mΩ). TTX was added to the bath solution to change concentrations in increments of 5–10 nM. Solutions were mixed and enough time allowed for equilibration. Traces were made at different times after change of TTX concentration as marked on the right of each trace. **a**, *nap^{ts}* larva. At 5 nM TTX, the e.j.p. is almost entirely blocked within 12 min. Synaptic transmission is not affected by TTX because graded potentials (bottom trace, superimposed recordings) can still be evoked electrotonically by strong stimuli (1.0 ms duration) even at 10 nM TTX when active conduction is blocked. **b**, Normal larva. The e.j.p. is lost only after 25 min in 40 nM TTX. At this point graded responses could be evoked electrotonically. The bath solution^{5,10} contains 0.8 mM Ca^{2+} .

‡ To whom correspondence should be addressed at University of Iowa.



about 4 ms in control larvae but is close to 10 ms in the mutant. This suggests that in *nap^{ts}* each axon contributing to the compound action potential has a longer than normal refractory period. Another less likely possibility is that *nap^{ts}* has a specific effect only on those axons with a short refractory period.

These effects of *nap^{ts}* on nerve conduction at permissive temperatures are unlikely to be due to a defect in a passive electrical property of excitable membranes such as increase in leakage conductance. Computer simulations of active conduction in a Hodgkin-Huxley model axon¹⁶ of the size found in *Drosophila* showed that the observed increase in refractory period and the conduction block at the restrictive temperature cannot be accounted for solely by an increase in leakage conductance (Joyner and Wu, unpublished data).

If *nap^{ts}* does affect the sodium channel, it will be of interest to identify other mutations with similar properties. Of eight other temperature-sensitive paralytic mutants on the X and second chromosomes that we recently examined, only *para^{ts}* shows behavioural and physiological phenotypes like those of *nap^{ts}*. Adult *para^{ts}* flies show rapid reversible paralysis at 29 °C (ref. 1). Simultaneous recording of the e.j.p. and nerve compound action potential indicates that the e.j.p. and compound action potential disappear at 37 °C but return when temperature is lowered (Fig. 3). In normal larvae, the e.j.p. and the compound action potential persist up to at least 40 °C (ref. 5). Synaptic transmission is not blocked in *para^{ts}* larvae because a graded e.j.p. could still be elicited electrotonically after failure of active conduction. Therefore, *para^{ts}*, like *nap^{ts}*, specifically blocks axonal conduction at high temperature, as originally suggested by Siddiqi¹⁷.

The *nap^{ts}* and *para^{ts}* gene products may share a related function because the double mutant is lethal even at temperatures permissive for each single mutant as shown in Table 1. A comparison of progeny classes 2 or 3 with 1 indicates that *nap^{ts}* or *para^{ts}* can cause a slight reduction in viability. However, the striking result is the complete failure to recover any double mutant progeny at all (class 4). The same is true for other *para* alleles in combination with *nap^{ts}* (Ganetzky and Wu, unpublished data). The lethality appears specific for the combination of *para* with *nap*, since other X-linked temperature-sensitive paralytic mutants show no such synergistic interaction with *nap^{ts}*.

In other examples of synthetic lethality in *Drosophila*, it has been possible to demonstrate at the biochemical level that the mutations involved have related functions¹⁸. If *nap^{ts}* is shown to

disrupt the assembly or function of sodium channels, a similar defect for *para^{ts}* might be indicated. However, the two mutants do not appear to act identically because at room temperature *para^{ts}*, unlike *nap^{ts}*, did not appear to differ significantly from wild type in TTX sensitivity or refractory period. One interesting possibility is that sodium channels are comprised of at least two different protein subunits—one altered by *nap^{ts}* and the other by *para^{ts}*.

Whether or not this notion proves correct, it is certainly the case that the gene products specified by the *nap* and *para* loci play essential roles in membrane excitability. By utilizing genetic and physiological techniques, it should be possible to identify other genes in *Drosophila* that control primary events of excitability in the nervous system. Moreover, we wish to emphasize that analysis of double mutants, as in the case of *nap^{ts}* and *para^{ts}*, can reveal functional relationships between gene products in excitable membranes that would be difficult to detect in any other way.

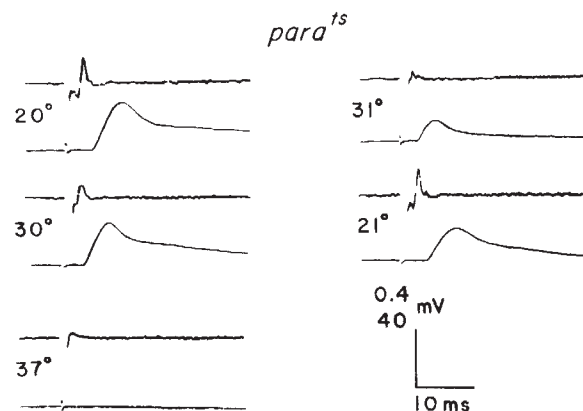


Fig. 3 Continuous recording of compound action potential and evoked e.j.p. from a *para^{ts}* larva at various temperatures. The e.j.p. (bottom trace in each panel) was recorded intracellularly from the same ventrolateral longitudinal muscle fibre. The compound action potential (top traces) was recorded from the nerve bundle containing the motor axons innervating the impaled muscle fibre. As the temperature is raised from 20° to 37 °C, both the e.j.p. and the compound action potential disappear (left). They reappear when the temperature is lowered back to 21 °C (right). The calcium concentration in bath solution was 0.8 mM.

Table 1 Progeny resulting from cross of
$$\frac{para^+ nap^{ts}}{Y nap^{ts}} \delta \delta \times \frac{para^{ts} Cy}{para^{ts} nap^{ts}} \varphi \varphi \text{ at } 23^\circ \text{C}$$

$\varphi \varphi$		$\delta \delta$	
$\frac{para^+ Cy}{para^{ts} nap^{ts}}$	$\frac{para^+ nap^{ts}}{para^{ts} nap^{ts}}$	$\frac{para^{ts} Cy}{Y nap^{ts}}$	$\frac{para^{ts} nap^{ts}}{Y nap^{ts}}$
2144	1270	1872	0

Cy (Curly) is a dominant wing marker located on a multiply inverted chromosome that does not undergo genetic exchange with its nap^{ts} bearing homologue. The Cy chromosome carries a normal allele of nap^{ts} . Four classes of progeny are expected to emerge from the above cross with equal frequencies. The four classes are distinguishable by sex and wing phenotype. Because both nap^{ts} and $para^{ts}$ are recessive, only the last class of progeny among the four is mutant for both $para^{ts}$ and nap^{ts} .

We thank S. Benzer for support and encouragement, L. Kauvar and R. Joyner for allowing us to cite work in progress and our colleagues especially M. Tanouye for comments on the manuscript. Work supported at various times by NSF grant PCM 73-01972 to S. B., Jane Coffin Childs Postdoctoral Fellowship and USPHS grant NS 15390 to B. G., and USPHS grant NS 15350 and NS 15797 to C.-F. W.

Received 28 May; accepted 16 June 1980.

1. Suzuki, D. T., Grigliatti, T. & Williamson, R. *Proc. natn. Acad. Sci. U.S.A.* **68**, 890–893 (1971).
2. Grigliatti, T. A., Hall, L., Rosenbluth, R. & Suzuki, D. *Mol. gen. Genet.* **120**, 107–114 (1973).
3. Ikeda, K., Ozawa, S. & Hagiwara, S. *Nature* **259**, 489–491 (1976).
4. Siddiqi, O. & Benzer, S. *Proc. natn. Acad. Sci. U.S.A.* **73**, 3253–3257 (1976).
5. Wu, C.-F., Ganetzky, B., Jan, L. Y., Jan, Y.-N. & Benzer, S. *Proc. natn. Acad. Sci. U.S.A.* **75**, 4047–4051 (1978).
6. Poody, C. A. & Edgar, L. *J. Cell Biol.* **81**, 520–527 (1979).
7. Gitschier, J., Strichartz, G. R. & Hall, L. M. *Biochem. biophys. Acta* **595**, 291–303 (1980).
8. West, G. J. & Catterall, W. A. *Proc. natn. Acad. Sci. U.S.A.* **76**, 4136–4140 (1979).
9. Narahashi, T. *Physiol. Rev.* **54**, 813–889 (1974).
10. Jan, L. Y. & Jan, Y.-N. *J. Physiol. Lond.* **262**, 189–214 (1976).
11. Katz, B. & Miledi, R. *Proc. R. Soc. Lond. B* **167**, 23–38 (1967).
12. Jan, Y.-N., Jan, L. Y. & Dennis, M. J. *Proc. R. Soc. Lond. B* **198**, 87–108 (1977).
13. Hardie, J. *J. Insect Physiol.* **22**, 661–668 (1976).
14. Hardie, J. & Osborne, M. P. *Comp. Biochem. Physiol.* **57A**, 59–66 (1977).
15. Kauvar, L. M. *CalTech Biol. ann. Report*, 154 (1979).
16. Hodgkin, A. L. & Huxley, A. F. *J. Physiol. Lond.* **117**, 500–544 (1952).
17. Siddiqi, O. in *Regulation of Growth and Differentiated Function in Eukaryotic Cells* (ed. Talwar, G. P.) 541–549 (Raven, New York, 1975).
18. Lucchesi, J. C. *Genetics* **59**, 37–44 (1968).

Non-uniform Ca^{2+} buffer distribution in a nerve cell body

D. Tillotson & A. L. F. Gorman

Department of Physiology, Boston University School of Medicine,
80 E. Concord Street, Boston, Massachusetts 02118

In nerve cells, Ca^{2+} influx through voltage-dependent channels in the membrane causes a transient rise in the intracellular, free Ca^{2+} concentration^{1–3}. Such changes have been shown to be important for the release of transmitter at the axon terminal^{3,4} and for the control of the movement of ions through channels in the soma membrane^{5,6}. The transient behaviour of the rise in Ca^{2+} concentration can, in part, be explained by the presence of sequestering systems in the cell which tend to limit the magnitude and duration of changes in internal Ca^{2+} (refs 7–10). It is possible that systems involved in buffering changes in internal Ca^{2+} are not distributed uniformly throughout the cell. This is particularly likely in the cell body, where a significant portion of the cytoplasm is occupied by the nucleus, whose buffering capacity may differ from that of other cellular regions. We report here that in the soma of a molluscan pacemaker neurone, the machinery responsible for short-term buffering of Ca^{2+} ions is localized near the inner surface of the plasma membrane.

The pacemaker neurone R-15 in the abdominal ganglion of *Aplysia californica* was used, and experiments were carried out in artificial seawater which contained (in mM) 500 NaCl, 10 KCl, 10 CaCl_2 , 50 MgCl_2 and 10 Tris at pH 7.7, and at a temperature of 16 °C. In some experiments, Na^+ was completely replaced by Tris. The procedures for iontophoretic injection of ions or pressure injection of purified arsenazo(III), voltage clamping the soma membrane and monitoring changes in intracellular Ca^{2+} with differential absorbance micro-spectrophotometry were similar to those described previously^{11,12}.

The experimental design is shown in Fig. 1a. Arsenazo(III) was injected and its intracellular concentration (about 200 μM) was determined using a small light-accepting probe (~150 μm diameter) which was about half the cell diameter. Starting near the inner membrane surface, an electrode containing 0.2 M CaCl_2 was advanced through the cytoplasm in 20- μm increments in a neurone clamped at a holding potential of about –40 mV, and absorbance changes associated with small test Ca^{2+} injections were measured at each increment. A fourth electrode containing 0.5 M EGTA was used to inject EGTA iontophoretically into the cell. The small light-accepting probe, used previously^{11,12} for measuring absorbance changes produced by Ca^{2+} injection from a microelectrode, has the disadvantage that

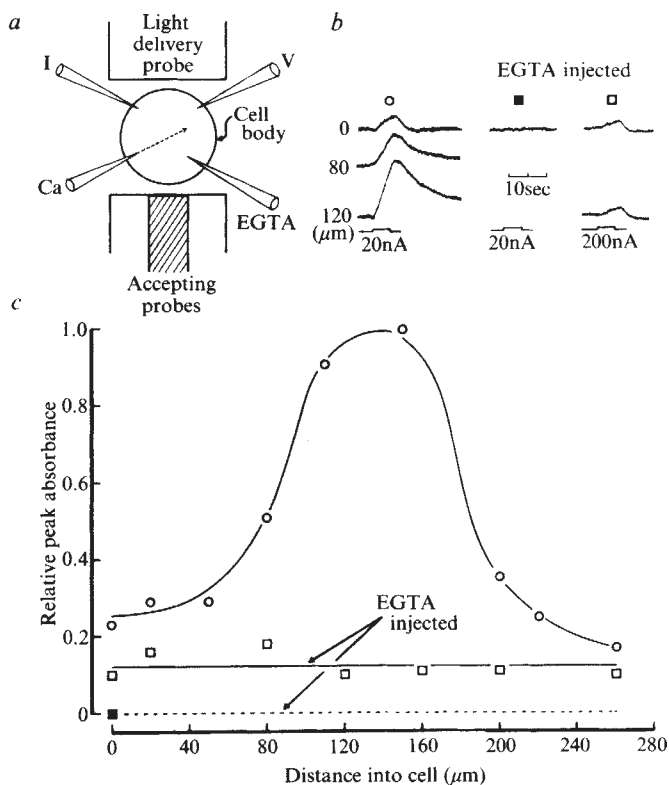


Fig. 1 Effects of Ca^{2+} injections at various distances in the soma cytoplasm on the free, intracellular Ca^{2+} concentration measured with arsenazo(III). *a* Shows the experimental arrangement. The nerve cell soma (cell body) was impaled with voltage (V), current (I), Ca^{2+} injection (Ca) and EGTA injection (EGTA) electrodes and positioned between a pair of fibre optic probes. Large and small (hatched lines) light-accepting probes are shown. *b* Shows the differential absorbance changes at 660–690 nm during Ca^{2+} injections for 5 s (20 or 200 nA) at the indicated distances into the cytoplasm from the cell membrane before (left column) and after injection of EGTA for 5 min (200 nA) (middle and right columns). The arsenazo(III) was 98.6% pure and its intracellular concentration was estimated to be about 200 μM . *c* Shows a plot of the normalized absorbance changes produced by Ca^{2+} injections (5 s, 20 nA) at various distances into the cell from the membrane before (open circles) and after (closed square) EGTA injection. The open squares plot the absorbance changes produced by a 10-fold increase in Ca^{2+} injection intensity (5 s, 200 nA) after EGTA injection. All data are from the same cell.

part of the cytoplasm and the plasma membrane are not directly in the light path. With all other factors equivalent, a test Ca^{2+} injection just outside the light path will produce a distorted absorbance signal, with slower rising kinetics and a reduced peak amplitude relative to those produced by an identical injection within the light path. To examine the localization of the Ca^{2+} buffer system, we replaced the small light-accepting probe (after injection of arsenazo(III)) with a probe of 400 μm diameter, which was larger than the average cell diameter ($\sim 300 \mu\text{m}$), thus rendering all points inside the cell equivalent. This system was tested experimentally using a 300- μm diameter droplet of solution (containing 300 mM KCl and 0.2 mM arsenazo(III) and buffered to a pH of 7.3 with 100 mM MOPS) in place of the neurone and injecting Ca^{2+} at different distances within the droplet. The changes in absorbance produced by Ca^{2+} injections were identical at all points within the droplet even at short times (within a few seconds of the injection), when elementary diffusion considerations dictate that the injected Ca^{2+} cannot have equilibrated throughout the droplet. This indicates that the sensitivity of the measurement system does not depend significantly on the position at which the injection is made. The use of a larger light probe, however, results in the transmission of a significant fraction of the incident light around the cell and hence severely limits our ability to quantify absorbance changes in terms of changes in internal Ca^{2+} concentration^{12,13}, although changes in the relative peak absorbance and in the kinetics of the absorbance can be determined.

Dye absorbance changes associated with small test Ca^{2+} injections (20 nA, 5 s) at different points in the cytoplasm are shown in the first column in Fig. 1b and a full set of measurements from the same cell are plotted in Fig. 1c. The amplitude of the absorbance signal is taken as an index of the buffering capacity of a region of the cytoplasm near the electrode tip. As the Ca^{2+} microelectrode was advanced into the centre of the cell, the amplitude of the absorbance signal increased and its kinetics decreased, indicating a diminution of buffering strength relative to the cell boundary, but as the electrode tip was moved towards the opposite membrane, the signal amplitude and kinetics returned to control values. Similar results were found in three other cells. The results from all four cells suggest that there may be a cytoplasmic layer near the cell membrane (0–50 μm) which has a different buffering capacity from that found in the centre of the cell.

To control for the possibility that the apparent Ca^{2+} buffering localization at the cell boundary is caused by a peculiarity of the absorbance recording arrangement, the Ca^{2+} chelating agent EGTA was injected iontophoretically into the cell for 5 min at 100 nA. An injection of this duration and intensity is more than sufficient to abolish all absorbance changes in cell R-15 produced by Ca^{2+} influx into the cytoplasm from the external medium during 0.5 s voltage-clamp pulses to all positive membrane potentials. We reasoned that a high internal EGTA concentration, if uniformly distributed throughout the cytoplasm, would render the whole cell equivalent with respect to buffering capacity, and that in these conditions Ca^{2+} injections at various distances through the soma would give absorbance changes of equal amplitude. The localization experiment was repeated after EGTA injection using a 10-fold greater injection current (200 nA, 5 s), because the previous injection level produced no measurable change in absorbance. The predicted result of equivalent intracellular buffering is shown in Fig. 1b and c. We conclude, therefore, that in normal circumstances the capacity to buffer Ca^{2+} is differentially distributed in the soma, being greatest in a cytoplasmic layer near the cell plasma membrane.

In most cells, buffering of intracellular Ca^{2+} is thought to be controlled by both sequestration by internal organelles and extrusion by systems in the plasma membrane^{7–9}. In nerve cells, a membrane-bound Na–Ca exchange system has been suggested as playing a significant part in the maintenance of intracellular Ca^{2+} levels¹⁴ and could explain the observed buffering localization, as Ca^{2+} near the membrane, and hence near the Na–Ca

exchange sites, would be extruded more quickly than Ca^{2+} in the cell interior. To test for this possibility, we examined the effect of removal of extracellular Na^+ on the absorbance change associated with Ca^{2+} injection near the inner membrane surface. Removal of external Na^+ should stop the exchange process^{1,14}, but it had no significant effect on the amplitude or time course of the absorbance change. This result agrees with previous findings from molluscan neurones¹⁵ and suggests that short-term Ca^{2+} buffering may not depend on the Na^+ -dependent transport of Ca^{2+} out of the cell. However, the situation with Ca^{2+} influx through the plasma membrane may be entirely different and membrane extrusion systems may be very much more important.

In nerve cells, changes in the total internal Ca^{2+} concentration are caused primarily by influx from the external medium, and evidence from squid axon suggests that the rise in free intracellular Ca^{2+} concentration is confined to the cell periphery^{1,16}. The localization of the machinery responsible for Ca^{2+} removal near the inner membrane surface may provide a physical basis for this confinement. It is well known that the nuclear membrane is freely permeable to large molecules and that the cytoplasm is continuous between the nuclear and non-nuclear regions of the cell; the localization of the buffer at the cell periphery therefore suggests that the buffering apparatus may be attached to some structural element in the cytoplasm. The identity of the systems responsible for this localization is uncertain, but could include membrane sites and cytoplasmic proteins which bind Ca^{2+} , and those intracellular organelles (such as mitochondria, endoplasmic reticulum) which sequester Ca^{2+} near the membrane.

This work was supported by USPHS grant NS11429.

Received 26 March; accepted 18 June 1980.

1. Baker, P. F., Hodgkin, A. L. & Ridgway, E. B. *J. Physiol., Lond.* **218**, 709–755 (1971).
2. Stinnakre, J. & Tauc, L. *Nature new Biol.* **242**, 113–115 (1973).
3. Llinás, R., Steinberg, I. Z. & Walton, K. *Proc. natn. Acad. Sci. U.S.A.* **73**, 2918–2922 (1976).
4. Katz, B. & Miledi, R. *J. Physiol., Lond.* **192**, 407–436 (1967).
5. Meech, R. W. A. *Rev. Biophys. Bioengng* **7**, 1–8 (1978).
6. Tillotson, D. *Proc. natn. Acad. Sci. U.S.A.* **76**, 1497–1500 (1979).
7. Baker, P. F. *Ann. N.Y. Acad. Sci.* **307**, 250–268 (1978).
8. Blaustein, M. P., Ratzlaff, R. W. & Kendrick, N. K. *Ann. N.Y. Acad. Sci.* **307**, 195–212 (1978).
9. Brinley, F. J. Jr. *Rev. Biophys. Bioengng* **7**, 363–392 (1978).
10. Meech, R. W. & Thomas, R. C. *J. Physiol., Lond.* **298**, 111–129 (1980).
11. Thomas, M. V. & Gorman, A. L. F. *Science* **196**, 531–533 (1977).
12. Gorman, A. L. F. & Thomas, M. V. *J. Physiol., Lond.* **275**, 357–376 (1978).
13. Thomas, M. V. *Biophys. J.* **25**, 541–548 (1979).
14. Blaustein, M. P. & Hodgkin, A. L. *J. Physiol., Lond.* **200**, 497–528 (1969).
15. Meech, R. W. & Thomas, R. C. *J. Physiol., Lond.* **265**, 867–879 (1977).
16. Mullins, L. J. & Requena, J. *J. gen. Physiol.* **74**, 393–413 (1979).

Ouabain stimulation of noradrenaline transport in guinea pig heart

Virendra K. Sharma, Lorelle A. Pottick & Shailesh P. Banerjee

Department of Pharmacology and Toxicology, University of Rochester School of Medicine and Dentistry, Rochester, New York 14642

Cardiac glycosides such as ouabain inhibit Na^+ -dependent high-affinity noradrenaline uptake in sympathetic nerve terminals by interacting with $(\text{Na}^+ + \text{K}^+)\text{ATPase}$ located in the neuronal membrane^{1–3}. It has been suggested that the augmentation of noradrenaline concentration at the neuromuscular junctions following the inhibition of its uptake may contribute to the ouabain-induced cardiac arrhythmia seen with toxic doses of the drug^{4,5}. Although it is generally accepted that the biochemical basis for the pharmacological response of cardiac glycosides must involve specific interaction of this class of drugs with the cardiac cell membrane Na^+ , K^+ -exchange pump, the precise molecular mechanism for the improvement of muscle performance by ouabain is controversial^{6,7}. Many investigators

claim that cardiac glycosides increase myocardial force of contraction by inhibiting $(\text{Na}^+ + \text{K}^+)\text{ATPase}$ ⁸⁻¹⁰. According to this hypothesis, such an inhibition might increase the intracellular concentration of Na^+ , which may augment the activity of the transmembrane Na^+ , Ca^{2+} -exchange pump described by several workers¹¹⁻¹⁴. We have found that ouabain stimulates Na^+ -dependent transport of noradrenaline by guinea pig heart sympathetic nerve terminals.

There is increasing evidence that low doses of cardiac glycosides increase the contractility of the myocardium and stimulate the active transmembranous transport of monovalent cations¹⁵⁻¹⁸. Such a conclusion has been vigorously challenged by Michael and co-workers^{19,20}. It is possible that the stimulation of cation transport by cardiac glycosides may be observed in intact cells as opposed to broken cell membrane preparations. Thus, Michael and his co-workers^{19,20}, who used an isolated purified enzyme system, were unable to observe any stimulation of $(\text{Na}^+ + \text{K}^+)\text{ATPase}$ activity by ouabain. We decided to investigate this possibility by examining the effects of different concentrations of ouabain on Na^+ -dependent noradrenaline transport in sympathetic nerve terminals of cardiac muscle. If the noradrenaline uptake system is tightly coupled to $(\text{Na}^+ + \text{K}^+)\text{ATPase}$ ¹⁻⁵, and low doses of ouabain can stimulate cation transport¹⁵⁻¹⁸, then these low concentrations of this cardiac glycoside must increase velocity of Na^+ -dependent noradrenaline uptake into the cardiac sympathetic nerve terminals.

The Na^+ -dependent [^3H]noradrenaline uptake in guinea pig heart slices was measured as described previously^{2,5}. The effect of several concentrations of ouabain on Na^+ -dependent [^3H]noradrenaline uptake is shown in Table 1. Although 1×10^{-10} M ouabain produced a small stimulation of the Na^+ -dependent [^3H]noradrenaline uptake, 1×10^{-9} M and 5×10^{-10} M increased the initial velocity of [^3H]noradrenaline transport into the cardiac sympathetic nerve terminals by about 28% and 63%, respectively. When the concentration of ouabain was raised to 5×10^{-9} M, a significant inhibition of uptake was

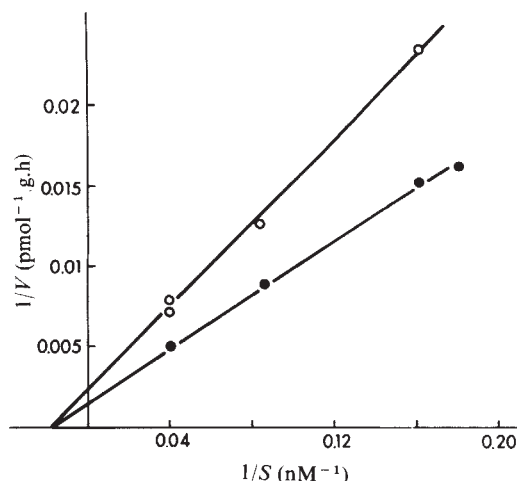


Fig. 1 Reciprocal plots of initial velocity of high-affinity noradrenaline uptake in guinea pig heart slices either in the presence (●) or absence (○) of 0.5 nM ouabain. Each point represents the mean of four determinations. Slopes and intercepts of the lines were determined by the method of least squares.

observed. Several experiments with different concentrations of ouabain (between 2×10^{-10} M and 8×10^{-10} M) indicated that maximal stimulation of Na^+ -dependent [^3H]noradrenaline uptake is seen at 5×10^{-10} M ouabain.

This increase of [^3H]noradrenaline uptake in sympathetic nerve terminals by ouabain may be due to an enhancement of the maximal velocity or apparent affinity of the uptake system for noradrenaline. The biochemical mechanism for the augmentation of [^3H]noradrenaline uptake in guinea pig slices by ouabain was evaluated by kinetic analysis. Results shown in Fig. 1 indicate that the maximal velocity ($V_{\max}$) for [^3H]noradrenaline uptake into the nerve endings of guinea pig heart slices is 444 pmol per g tissue per h, with the apparent affinity constant (K_m) for noradrenaline being 5.6×10^{-8} M. The presence of 5×10^{-10} M ouabain increased the maximal velocity for [^3H]noradrenaline uptake to 667 pmol per g tissue per h without altering the K_m for the neurotransmitter (Fig. 1). These results indicate that ouabain may stimulate [^3H]noradrenaline uptake in cardiac sympathetic nerve terminals by increasing the $V_{\max}$ without affecting K_m .

There are several implications arising from this observation of stimulation of Na^+ -dependent noradrenaline transport in guinea pig heart sympathetic nerve terminals by ouabain. First, there is some controversy on the involvement of $(\text{Na}^+ + \text{K}^+)\text{ATPase}$ in the transport for noradrenaline²¹. Recent evidence from several laboratories indicates an intimate relationship between these two transport systems¹⁻⁵. For example, ouabain has been found to show identical species sensitivity for the inhibition of $(\text{Na}^+ + \text{K}^+)\text{ATPase}$ and noradrenaline uptake in heart and spleen slices¹⁻³. Again, in regenerating sympathetic nerve terminals, ouabain binding sites appear before noradrenaline uptake systems can be detected⁵. As the low concentration of ouabain that is known to stimulate the Na^+ , K^+ -exchange pump¹⁵⁻¹⁸ activates noradrenaline uptake (Table 1), these two transport systems must be coupled. Second, there is some dispute regarding the mechanism by which ouabain alters the level of noradrenaline at the neuromuscular junction. It has been suggested that ouabain may stimulate release of noradrenaline from sympathetic nerve terminals without affecting the uptake mechanism^{22,23}. As ouabain cannot stimulate noradrenaline uptake by augmenting the release of accumulated noradrenaline, there has to be some direct effect of ouabain on high-affinity noradrenaline uptake mechanisms.

In conclusion, this report describes for the first time a stimulation of Na^+ -dependent noradrenaline uptake in myocardial sympathetic nerve terminals by ouabain. Since ouabain selectively interacts with cell surface $(\text{Na}^+ + \text{K}^+)\text{ATPase}$, the results

Table 1 Effect of different concentrations of ouabain on high-affinity noradrenaline uptake in guinea pig heart slices

Concentration of ouabain (pM)	Specific [^3H]noradrenaline uptake (pmol g ⁻¹ h ⁻¹)*
0	83.8 ± 2
1	77.1 ± 1
50	79.1 ± 6
100	96.6 ± 7
500	136.7 ± 3†
1000	107.8 ± 6†
5000	37.2 ± 4‡

Male guinea pigs were killed by a blow on the head. Ventricle of the heart was removed immediately and placed in modified Krebs solution (4 °C) and cleared of extraneous tissues as described previously. Modified Krebs solution had the following composition (mM): NaCl, 118; KCl, 4.7; CaCl₂, 2.5; MgCl₂, 0.54; NaHCO₃, 25; NaH₂PO₄, 1; and glucose, 11. EDTA (20 μg ml⁻¹), ascorbic acid (200 μg ml⁻¹) and nialamide (1.25×10^{-5} M), were added to retard the spontaneous oxidation of noradrenaline. Ventricle slices were prepared with a Stadie-Riggs tissue microtome. The slices were weighed and immersed in 10 ml beakers containing 3.8 ml of Krebs solution. All data were obtained from triplicate slices and each experiment was repeated at least once more. Varying concentrations of ouabain (0.1 ml volume) and 26 nM [^3H]noradrenaline were added and the slices were shaken in a Dubnoff metabolic incubator at 37 °C in an atmosphere of 95% O₂ and 5% CO₂ for 10 min. Preliminary experiments indicated that the rate of [^3H]noradrenaline uptake remained linear for at least 10 min when the weight of the slices was between 20 and 40 mg. Nonspecific accumulation of radioactivity or blanks was routinely determined in triplicate by parallel incubation at 4 °C for 10 min. After incubation, slices were rinsed three times with ice-cold Krebs solution, blotted on filter paper, and placed in 1 ml of ice-cold 0.4 N perchloric acid. Tissues were extracted in the acid for 21 h at 4 °C and this was followed by centrifugation at 3,000g for 10 min in a Sorvall centrifuge (Model GLC-2). An aliquot (0.5 ml) of each sample was transferred to a vial containing 9.5 ml of Scintiverse (Fisher) and counted in a Packard Tri-Carb liquid scintillation spectrometer (Model 3380) at 30% efficiency as determined with internal standards. Tissue/medium ratios for [^3H]noradrenaline uptake were calculated as (c.p.m. per g wet weight tissue)/(c.p.m. per ml of medium). Results shown are averages of four experiments and each experiment was run in triplicate.

* Mean ± s.e.m.

† $P < 0.01$ compared with control.

‡ $P < 0.001$ compared with control.

of this report provide further insight into the relationship between the monovalent cation transport pump and the noradrenaline uptake system.

This work was supported in part by grant HL-18185 from NIH.

Received 17 January; accepted 18 June 1980.

- Sharma, V. K. & Banerjee, S. P. *Fedn Proc.* **35**, 405A (1976).
- Sharma, V. K. & Banerjee, S. P. *Eur. J. Pharmac.* **41**, 417-429 (1977).
- Stickney, J. L. *Res. Commun. chem. Path. Pharmac.* **14**, 227-235 (1976).
- Sharma, V. K. & Banerjee, S. P. *Molec. Pharmac.* **13**, 796-804 (1977).
- Sharma, V. K. & Banerjee, S. P. *Molec. Pharmac.* **15**, 35-42 (1979).
- Schwartz, A., Lindenmayer, G. & Allen, J. C. *Pharmac. Rev.* **27**, 3-134 (1975).
- Godfraind, T. *Biochem. Pharmac.* **24**, 823-827 (1975).
- Besch, H. R., Jr., Allen, J. C., Glick, G. & Schwartz, A. *J. Pharmac. exp. Ther.* **171**, 1-12 (1970).
- Akera, T., Larsen, F. S. & Brody, T. M. *J. Pharmac. exp. Ther.* **173**, 145-151 (1970).
- Langer, G. A. *Circulation* **46**, 180-187 (1972).
- Baker, P. F. *Prog. biophys. mol. Biol.* **24**, 177-223 (1972).
- Langer, G. A. *Rev. Physiol.* **35**, 55-86 (1973).
- Reuter, H. *Circulation Res.* **34**, 599-605 (1974).
- Brading, A. F. *Phil. Trans. R. Soc. B* **265**, 35-46 (1973).
- Ghysel-Burton, J. & Godfraind, T. *Br. J. Pharmac.* **66**, 175-184 (1979).
- Cohen, I., Daut, J. & Noble, D. *J. Physiol., Lond.* **260**, 55-74 (1976).
- Blood, B. E. *J. Physiol., Lond.* **251**, 69P (1975).
- Blood, B. E. & Noble, D. *J. Physiol., Lond.* **266**, 76A (1977).
- Michael, L., Pitts, B. J. R. & Schwartz, A. *Science* **200**, 1287-1289 (1978).
- Michael, L. H., Schwartz, A. & Wallick, E. T. *Molec. Pharmac.* **16**, 135-146 (1979).
- White, T. D. & Keen, P. *Molec. Pharmac.* **7**, 40-45 (1971).
- Garcia, A. G. & Kirpekar, S. M. *Br. J. Pharmac.* **47**, 729-747 (1973).
- Paton, D. M. *Br. J. Pharmac.* **49**, 614-627 (1973).

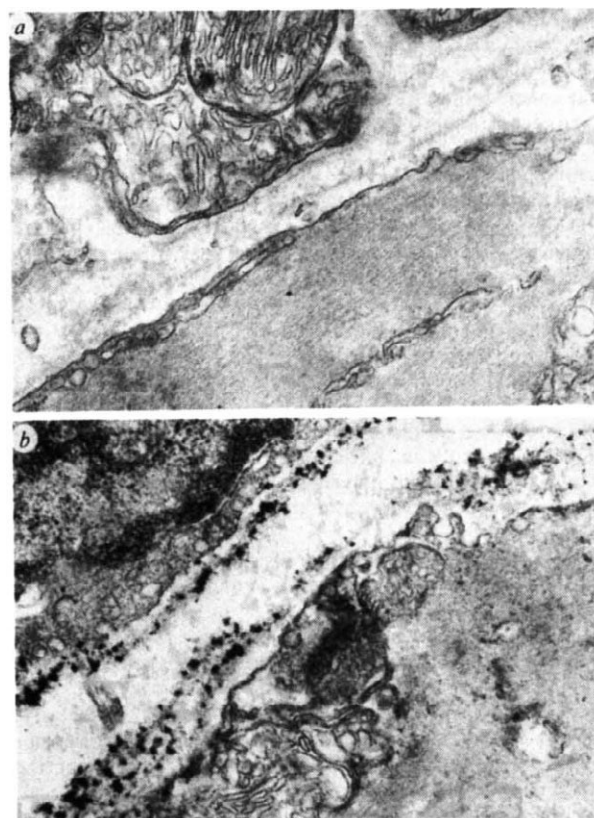


Fig. 1 Electron micrographs of neuraminidase-treated (a) and control (b) guinea pig left atria. Fixation and lanthanum staining were as described in the text. Note the heavy deposits of lanthanum on the external matrix of control cells, and their absence in neuraminidase-treated atria. $\times 30,000$.

Removal of sialic acid from cardiac sarcolemma does not affect contractile function in electrically stimulated guinea pig left atria

S. E. Harding & J. Halliday

Department of Pharmacology, King's College, London WC2R 2LS, UK

Calcium is essential for the maintenance of contraction in the heart, and it has been suggested that the glycoprotein matrix on the external surface of cardiac cells is a critical factor in the supply of 'activator' calcium to the beating heart¹. Sialic acid residues are important calcium-binding sites in the matrix and treatment with neuraminidase, an enzyme which cleaves sialic acid from oligosaccharide chains, has been reported to abolish spontaneous contraction in cultured heart cells, without causing general breakdown in the plasma membrane^{2,3}. However, we report here that in intact, electrically stimulated guinea pig atrial preparations, neuraminidase treatment produces no significant changes in resting tension or force of contraction. It is also without effect on the response to calcium removal and replacement, and inotropic or toxic concentrations of cardiac glycosides. In these experiments, up to 79% of the total tissue sialic acid was removed, and electron microscopy studies showed that this removal was accompanied by marked reduction of lanthanum binding to the sarcolemma. We therefore conclude that the calcium-binding sialic acid residues in the external matrix of the cardiac cells are of little importance in the maintenance of contractile function in these intact atrial preparations.

In preliminary determinations of the total sialic acid content of guinea pig heart, by incubation of the tissue in 0.05 M H_2SO_4 at 80 °C for 1 h (a modification of the method of Bers⁴), the content in nmol sialic acid per mg wet weight was 1.64 ± 0.12 for atria ($n=6$) and 1.49 ± 0.14 (s.e.m.) for ventricles ($n=4$). Although these values are not significantly different, the time course of sialic acid release by neuraminidase was found to be much slower in ventricles than in atria. By method (1) (see Table 1 legend), neuraminidase treatment freed 79% of total sialic acid from the atria within 5 h, whereas the corresponding release

from ventricles was only 24%. We therefore decided to use atria in all further experiments.

For the experiments in which the force of contraction was monitored, two methods were used for the incubation of the tissue with neuraminidase (see Table 1 legend). The first, in which the atrial strip was detached from the electrodes and incubated in a concentrated enzyme solution, had the advantage that all the neuraminidase-sensitive sialic acid, 1.30 ± 0.06 nmol per mg wet weight ($n=5$), was removed in the 5-h period (the mean percentage difference between the amount released at 5 h and that at 6 h was $2.4 \pm 1.5\%$, and was not significant). This lack of further release was not due to deterioration of the enzyme, because the 5-h incubation medium was as effective as fresh neuraminidase solution in releasing sialic acid from *N*-acetyl-neuramin-lactose (from bovine colostrum, Sigma). The second method, in which the enzyme was added to the Krebs solution in the organ bath, enabled the force of contraction and resting tension to be monitored throughout the incubation. However, the sialic acid content of the bath could not be measured directly with this method because the volume of dilution was too great; parallel incubations in small volumes, but with the same ratio of enzyme to Krebs, were required to estimate the release. The amount released by this procedure was 1.07 ± 0.04 nmol sialic acid per mg wet weight ($n=5$). This represents 65% of the total (82% of the neuraminidase-sensitive) sialic acid content, which is less than that achieved with method (1), but still higher than the 61% removal shown to cause disruption of the external matrix of cultured heart cells³.

Electron microscopy studies were done to confirm visually that neuraminidase treatment was removing sialic acid residues from the external matrix. Tissues were exposed to concentrated neuraminidase (1 unit ml^{-1}) for 5 h, then transferred to HEPES buffer, composition (mM) 133 NaCl, 3.6 KCl, 1.0 $CaCl_2$, 0.3

MgCl₂, 16.0 glucose, 3.0 HEPES (pH 7.1), containing 5 mM LaCl₃ for 2 h. The temperature was maintained at 37 °C and the solutions were continuously bubbled with 100% O₂. Control tissues were exposed to the acetate buffer alone (see Table 1 legend) during the first 5 h. The tissues were fixed as described in ref. 3.

Preliminary results showed that in neuraminidase-treated tissues, lanthanum binding is greatly reduced in all areas (Fig. 1a). Control tissues, however, demonstrate the normal pattern of lanthanum binding described by other workers³; that is, large accumulations of lanthanum at the external matrix of the sarcolemma (Fig. 1b). This finding supports the biochemical data suggesting that neuraminidase treatment of whole atria is effective in removing at least part of the glycocalyx.

As may be seen in Table 1, the force of contraction of the control atrial preparations declines over the time course of these experiments. This is a well reported phenomenon, termed 'ageing', which Carrier *et al.* have suggested⁵ is due to a decrease in calcium binding at the external surface of the sarcolemma; however, sialic acid removal neither enhanced nor attenuated this fall in force, casting some doubt on this hypothesis. In addition, sialic acid removal did not alter the resting tension, indicating that the capacity of the heart to relax was unchanged, and implying that the systems involved in the maintenance of low intracellular calcium between beats had not been affected.

We were particularly interested in the effect of neuraminidase treatment on the response to ouabain, as several reports have suggested that positive inotropic concentrations of cardiac glycosides increase external calcium binding. Ouabain, for example, increased both the density of the negative charges on the cardiac membrane⁶ and the amount of La³⁺-displaceable calcium bound to this surface⁷. We found, however, that neuraminidase treatment had no effect on either the positive inotropic response to 0.2 µM ouabain or the maximum rise in force evoked by a higher concentration. The toxicity of ouabain was assessed by the time to onset of arrhythmias after administration of a toxic concentration, and the rise in resting tension 30 min after this administration. Neither parameter was altered by sialic acid removal (see Table 1).

In the experiments with cultured heart cells, it was reported that neuraminidase treatment increased the amount of ⁴⁵Ca taken up², and the speed of uptake and washout⁸. This observation suggested that the external matrix might in some way regulate calcium entry to or exit from the cell. We therefore examined the effect of neuraminidase on the rate of fall in force of atrial preparations after removal of calcium from the bathing

Table 2 Effects of neuraminidase treatment on rates of decrease or recovery of force with calcium removal or replacement, respectively

	Half-time (s) for decrease in force of contraction following Ca ²⁺ removal		Half-time (s) for increase in force of contraction following Ca ²⁺ replacement	
	Before 5 h incubation	After 5 h incubation	Before 5 h incubation	After 5 h incubation
Control (n = 6)	22.9 ± 2.7	16.4 ± 1.1	45.5 ± 3.1	57.5 ± 6.9
Neuraminidase-treated (n = 6)	21.3 ± 2.2	18.2 ± 1.1	57.4 ± 4.9	58.3 ± 5.8

Method (2) (see Table 1 legend) was used for neuraminidase treatment of the atrial strip. For calcium removal, the organ bath was drained, then washed through with prewarmed, pre-oxygenated, calcium-free Krebs (calcium not osmotically replaced). Half-time for the decrease in force of contraction was taken as the time for the force to decrease to 50% of that immediately before draining the bath. After 2 min, calcium was re-introduced by injection of 100 mM CaCl₂ solution to give a bath concentration of 2.5 mM. Half-time for recovery was the time to reach ½(max-min) force, where max = maximum force of contraction attained within 10 min of calcium replacement, and min = force immediately before replacement. This procedure was repeated three times before and three times after neuraminidase treatment. All values are mean ± s.e.m.

medium, and the subsequent rise on calcium replacement. The method, and the procedure for calculating half-times, are given in Table 2 legend. No differences in rates of decline and recovery of force of contraction between control and neuraminidase-treated preparations could be detected; a slight decrease was seen in the time taken for the force to fall, but the control also showed this change. It was possible to resolve the fall in force after calcium removal into two exponential decay processes⁹, and therefore, in case there had been opposing changes in the two processes that were not apparent in the overall half-times, each curve was analysed, and the intercept on the ordinate and the half-life calculated for both the fast and slow decays. Once again, there was no significant difference between control and neuraminidase-treated preparations, indicating that there had been no change in the rate of loss of calcium from, or return of calcium to, those pools associated with maintenance of contraction.

In our experiments, therefore, removal of a large proportion of sialic acid from the membrane had no detectable effect on the function of the intact myocardium. These results are difficult to reconcile with those obtained in cultured cells, where spontaneous beating ceased and calcium exchangeability rose after neuraminidase treatment². A possible explanation for this difference is that the culturing of the cells, which includes trypsin

Table 1 Effect of neuraminidase treatment on resting tension and force of contraction and response to ouabain in guinea pig atria

	Force of contraction after incubation (%)*	Decrease in resting tension during incubation (%)*	Response to ouabain			
			Increase in force of contraction with 0.2 µM ouabain (%)†	Maximum increase in force with ouabain (%)†	Time to onset of arrhythmias after addition of 1 µM ouabain (min)	Increase in resting tension with 1 µM ouabain (%)†
Method (1) (n = 4)						
Control	13.5 ± 7.1	—	173 ± 10.3	218 ± 37.0	—	—
Neuraminidase-treated	18.5 ± 5.3	—	163 ± 22.1	292 ± 19.8	—	—
Method (2) (n = 8)						
Control	41.9 ± 4.3	28.5 ± 14.2	157 ± 7.2	204 ± 8.9	11.8 ± 4.2	7.5 ± 3.4
Neuraminidase-treated	46.8 ± 2.7	43.0 ± 14.2	153 ± 5.5	211 ± 11.6	18.8 ± 4.5	10.2 ± 6.1

Left atria were halved and mounted in contact with platinum electrodes in a 10-ml organ bath containing Krebs (119 mM NaCl, 4.7 mM KCl, 0.94 mM MgSO₄·7 H₂O, 1.2 mM KH₂PO₄, 25 mM NaHCO₃, 11 mM glucose, 2.5 mM CaCl₂) bubbled with 95% O₂/5% CO₂, and stimulated at 3 Hz, 500 µs, supramaximal voltage. In method (1) the tissue was then unhooked from the electrode and placed in 1 ml acetate buffer (50 mM Na acetate, 154 mM NaCl, 9 mM CaCl₂, 5 mM KOH and 100 mM Tris-HCl, pH 7.4) containing 1 unit ml⁻¹ neuraminidase (Behringwerke: 1 unit will split off 1 µmol N-acetylneuraminic acid from human α₁-acid glycoprotein in 1 min at 37 °C and pH 5.5), and bubbled with 100% O₂. Aliquots (0.2 ml) of the incubation medium were assayed for free sialic acid by the Warren method¹⁰. After 5 h, when all the neuraminidase-sensitive sialic acid had been split off, the preparation was removed, resuspended in the organ bath and stimulation restarted. After a period of equilibration, ouabain was added to the bath to a concentration of 0.2 µM and the positive inotropic effect allowed to develop for 30 min. In method (2) neuraminidase was added directly to the organ bath to give a concentration of 0.1 units ml⁻¹, and the tissue kept beating throughout the incubation. Parallel experiments using smaller volumes were done to estimate release of sialic acid by this method. After 5 h the neuraminidase was removed by washing the tissue six times with Krebs solution; ouabain was then added to a bath concentration of 0.2 µM as before. After 30 min the concentration was raised to 1 µM and the force of contraction and resting tension were monitored for a further 30 min. The time to onset of arrhythmias and increase in resting tension were taken as indices of ouabain toxicity. Control tissues were treated similarly to test ones, except that neuraminidase was omitted from the acetate buffer. All values given are the mean ± s.e.m. No neuraminidase-treated values differ significantly from control (P < 0.05).

* Expressed as a percentage of that immediately before incubation with neuraminidase.

† Expressed as a percentage of that immediately before addition of ouabain.

digestion, may have in some way increased the susceptibility of the matrix to disruption by sialic acid removal. Alternatively, it is just possible that similar changes in calcium permeability did occur in our preparations, but that the reserve capacity of the calcium-sequestering systems of the cell was sufficient to compensate for the extra influx.

The cessation of beating in the neuraminidase-treated, cultured cells might not be due to failure of contraction mechanisms as such, but to a disturbance of pacemaker mechanisms; this could explain why our preparations, which were electrically stimulated, continued to beat after a similar treatment. However, preliminary experiments with spontaneously beating right atria have shown that contractile activity is maintained throughout 5 h incubation in concentrated neuraminidase (1 unit ml⁻¹) at 37 °C. It therefore seems unlikely that sialic acid removal results in a disruption of pacemaker mechanisms.

We have thus shown that treatment of guinea pig atria with neuraminidase can remove up to 79% of tissue sialic acid and strip lanthanum-binding sites from the external sarcolemma. This disruption of the external glycocalyx, however, has no detectable effect on the force of contraction of electrically stimulated atria or the response of these atria to ouabain or calcium-free media. It seems, therefore, that the contribution of the sialic acid residues in the external glycoprotein matrix to contractile function in the guinea pig atrium is negligible.

We wish to thank Mr G. Morgan and Miss K. Bugg for the electron microscope studies, and Drs I. K. M. Morton and J. Littleton for their helpful discussion. S.H. is an MRC student.

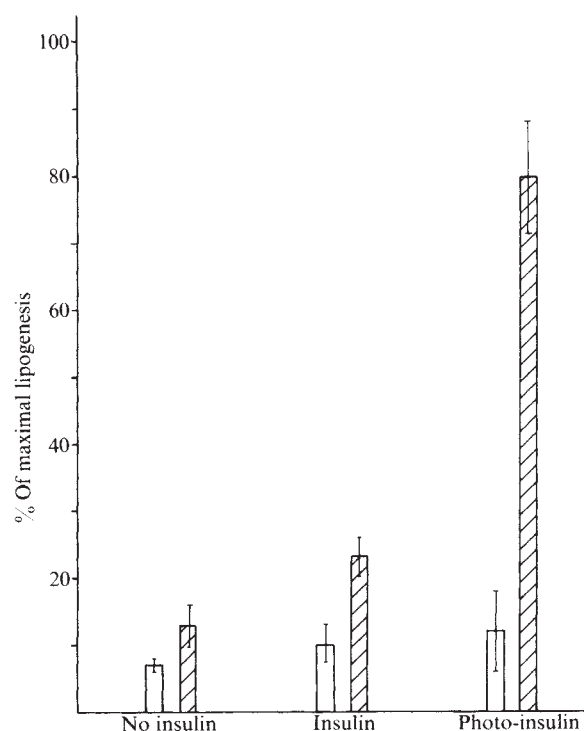


Fig. 1 Effect of covalent binding of photolabelled insulin on lipogenesis in adipocytes. Experiments were carried out with irradiation (▨) and without irradiation (□) with the indicated preincubations. Per cent values (height of columns) are the ratios of basal lipogenesis (c.p.m. without added insulin) to maximal response (c.p.m. at 10 ng insulin per ml) and represent averages from three independent experiments. Adipocytes were prepared¹³ from epididymal fat tissue of four Sprague-Dawley rats (130–160 g) and resuspended in 90 ml of Krebs-Ringer-HEPES buffer, pH 7.4, containing 1% bovine serum albumin¹⁴. Three lots of 25 ml (2–4 × 10⁵ cells per ml) were transferred to thermostatted (37 °C) Millipore ultrafiltration chambers (47 mm internal diameter) equipped with boiled Millipore-MF-filters (8 μm pore size) and Teflon-coated stirrer bars. All subsequent operations up to the irradiation period were carried out in light of long wavelength (λ 500 nm). No addition was made to the first chamber (blank). Insulin (monocomponent pork, Novo; 14 pmol ml⁻¹) was added to the second chamber and Napa-DP-insulin (14 pmol ml⁻¹) to the third. The suspensions were equilibrated under slow (~100 r.p.m.) stirring for 5 min. Each chamber was then irradiated from above for 1 min with a Philips HPK 125 W/L UV lamp equipped with a black glass filter (UVW-55, Hanau Wertheim). The lamp was 8 cm from the surface of the stirred adipocyte suspension. After irradiation, the Millipore chambers were sealed. Warm (37 °C) Krebs-Ringer-HEPES buffer (300 ml) was run through in a constant volume washing procedure at a flow rate of 10 ml min⁻¹. After washing, 1-ml aliquots (quadruplicates) of the suspensions were used in a lipogenesis assay according to ref. 8. The incubation time was 2 h. Control experiments were carried out as above, but without irradiation.

Covalent linking of photoreactive insulin to adipocytes produces a prolonged signal

Dietrich Brandenburg, Cornelia Diaconescu, Derek Saunders & Peter Thamm

Deutsches Wollforschungsinstitut, Veltmanplatz 8, D-5100 Aachen, FRG

The first step of insulin's many cellular functions is specific binding to receptors on the plasma membrane of target cells. The subsequent molecular basis of insulin action, particularly the coupling mechanism(s) involved in transmitting the biological message, remains largely unknown¹. Our approach to the problem centres on the application of a series of well characterized photo-insulins carrying an aryl-azido or nitro-aryl-azido group in positions A1, B1, B2 or B29 (refs 2–4). Specific binding to membrane components could be demonstrated with B1-, B29- (ref. 5) and A1-photo-insulins⁶ as well as a B2-derivative⁷. We now report that lipogenesis is increased to, and maintained at, near-maximal levels for several hours after photoinduced covalent binding of B2- (2-nitro,4-azidophenyl-acetyl)-des-Phe^{B1}-insulin (Napa-DP-insulin⁸) to living adipocytes.

A fat cell assay⁸ was used to determine the biological effects of Napa-DP-insulin. Adipocytes were first incubated with Napa-DP-insulin (80 ng ml⁻¹) to obtain a receptor occupancy of more than 90% (ref. 9). The photoreactive group was activated by short exposure to UV light of wavelengths 300–400 nm. Cells were then washed for 30 min by a mild constant-volume procedure which diluted the insulin concentration by 1 : 3 × 10⁴. The wash period was sufficient to allow for dissociation of the non-covalently bound insulins from the receptor. These adipocytes were then used in a standard lipogenesis assay to determine the effect of prebound and added insulin. In control experiments, adipocytes were incubated in the dark or under UV light with or without unmodified insulin.

It was important to find conditions of activation which would not damage the adipocytes. Appropriate experiments in which

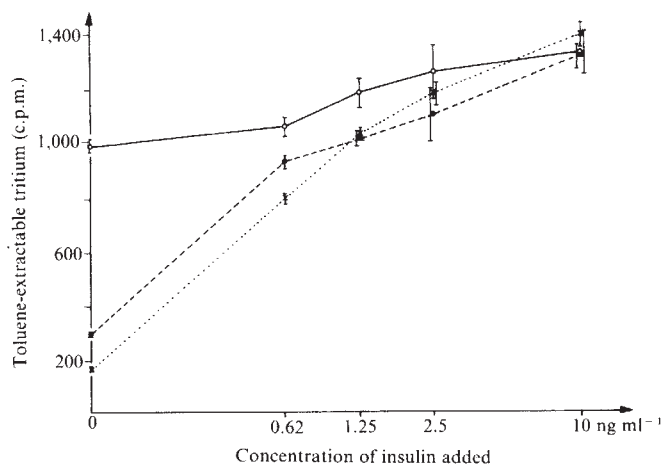


Fig. 2 Insulin-stimulated incorporation of tritium from D-[2-³H] glucose into toluene-extractable lipids of isolated rat fat cells after the following pretreatment: curve 1 (·····), irradiation only; curve 2 (-----), preincubation with insulin (14 pmol ml⁻¹) + irradiation; curve 3 (—), preincubation with Napa-DP-insulin (14 pmol ml⁻¹) + irradiation.

adipocytes were irradiated for various periods showed that there was no appreciable loss in lipogenic viability after 1 min of irradiation. This represents twice the half life of the photo-label in the same conditions, an activation of about 75% (ref. 4).

Adipocytes preincubated with high concentrations of insulin and subsequently washed gave basal values which were only slightly higher than normal (Fig. 1), and to a somewhat greater extent in the case of Napa-DP-insulin. Adipocytes responded normally to increasing concentrations of insulin in the subsequent lipogenesis assay (data not shown). The higher basal values may be related to the memory effect described by Gliemann¹⁰.

Irradiation of adipocytes which had been preincubated with insulin caused basal lipogenesis to increase from 11 to 23% (Fig. 1). However, the concentration dependence and maximal stimulation in the lipogenesis assay were similar to those of cells preincubated without insulin (see curves 1 and 2 in Fig. 2). A considerable increase in basal lipogenesis was observed after irradiation of adipocytes preincubated with photoreactive insulin (Fig. 1 and Fig. 2, curve 3). This amounted to 75% of the maximal stimulation in two experiments and to 90% in a third. The difference between the experiments was probably due to small uncontrollable variations in the conditions of irradiation.

The addition of progressively higher concentrations of insulin to the medium resulted in a further incorporation of tritium into lipids until maximal response was reached. In all three experiments, the maximal stimulation of lipogenesis in adipocytes preincubated with or without insulin or Napa-DP-insulin was identical within experimental error. Maintenance of normal insulin sensitivity in the latter case indicates that the adipocytes were still fully viable and that the effects were not due to specific damage through photo-sensitization.

Our study demonstrates that photoreactive insulin can be activated without interfering with the viability of isolated adipocytes. Furthermore, covalent binding indeed generates a biological response. This response lasts for a considerable period. The photo-insulin-treated cells are handled for a period of 0.75 h (washing and pipetting) before the 2-h incubation with ³H-glucose, so that the covalent hormone-receptor complex must be functional over at least 2.75 h. However, it is not inconceivable that the complexes are functional as long as they are intact.

Specific covalent binding of insulins with photoreactive groups at the N-terminus of the B-chain to cell membrane preparations has been demonstrated^{6,7,11}. Therefore, we deduce

that the effect we observed was due to specific linkage to the adipocyte receptor. It is interesting that the N-terminus of the B-chain has previously not been included in the putative receptor-binding region of the hormone¹². The irreversible complex that is formed is obviously able to elicit the subsequent events of glucose transport and metabolism. Experiments are under way to determine whether a given population of covalently bound hormone-receptor complexes has the same ability to stimulate lipogenesis and glucose transport as the same population of noncovalently linked hormone-receptor complexes.

This work was supported by Sonderforschungsbereich 113 Diabetesforschung Düsseldorf.

Received 7 March; accepted 24 June 1980.

1. Czech, M. P. A. *Rev. Biochem.* **46**, 359-384 (1977).
2. Thamm, P. & Brandenburg, D. in *Peptides 1978—Proc. 15th Eur. Peptide Symp.*, Poland (eds Siemion, Z. & Kupryszewski, G.) 639-642 (Wrocław University Press, Poland, 1979).
3. Thamm, P., Saunders, D. & Brandenburg, D. in *Proc. 2nd int. Insulin Symp.*, Aachen (eds Brandenburg, D. & Wollmer, A.) 309-316 (De Gruyter, Berlin, 1980).
4. Saunders, D., Thamm, P. & Brandenburg, D. in *Proc. 2nd int. Insulin Symp.*, Aachen (eds Brandenburg, D. & Wollmer, A.) 317-325 (De Gruyter, Berlin, 1980).
5. Wisher, M. H. *et al. Diabetologia* **15**, 282 (1978).
6. Wisher, M. H. *et al. Biochem. biophys. Res. Commun.* **92**, 492-498 (1980).
7. Kühn, L., Meyer, H., Rutschmann, M. & Thamm, P. *FEBS Lett.* **113**, 189-192 (1980).
8. Moody, A. J., Stan, M. A., Stan, M. & Gliemann, J. *Horm. Metab. Res.* **6**, 12-16 (1974).
9. Gammeltoft, S. & Gliemann, J. *Biochim. biophys. Acta* **320**, 16-32 (1973).
10. Gliemann, J., Gammeltoft, S. & Vinten, J. *J. Biol. Chem.* **250**, 3368-3374 (1975).
11. Yip, C. C., Yeung, C. W. T. & Moule, M. L., *Biochemistry* **19**, 70-76 (1980).
12. Pullen, R. A. *et al. Nature* **259**, 369-373 (1976).
13. Rodbell, M. *J. Biol. Chem.* **239**, 375-380 (1964).
14. Gliemann, J. & Vinten, J. *J. Physiol., Lond.* **236**, 499-516 (1974).

A monomeric insulin from the porcupine (*Hystrix cristata*), an Old World hystricomorph

R. Horuk*, T. L. Blundell*, N. R. Lazarus†, R. W. J. Neville†, D. Stone‡ & A. Wollmer§

* Laboratory of Molecular Biology, Department of Crystallography,

† Biochemical Research Department, The Wellcome Foundation Ltd.,

‡ Biochemical Research Department, The Wellcome Foundation Ltd., Temple Hill, Dartford, Kent DA1 5AH, UK

§ Protein Chemistry R&D, Wellcome Research Laboratories, Langley Court, Beckenham, Kent, UK

§ Fachgebiet Struktur und Funktion der Proteine, Abteilung Physiologische Chemie, Rheinisch-Westfälische Technische Hochschule, D-5100 Aachen, FRG

The insulins of New World hystricomorph rodents exhibit many novel amino acid changes in primary structure when compared with other mammalian insulins¹⁻³. These changes give rise to unusual properties (low potency, failure to self-associate⁴⁻⁶) not shared by other naturally-occurring insulins. We report here on the primary structure, zinc-binding properties and circular dichroism (CD) of porcupine insulin (*Hystrix cristata*), the first Old World hystricomorph insulin to be investigated, and discuss the changes in primary structure of the hormone in relation to its properties. Residue B₂₂ is strongly implicated as being responsible for the unusual properties of porcupine insulin.

Porcupine insulin differs at seven residues from bovine insulin (Fig. 1) and is unable to bind zinc (Fig. 2). Such behaviour has been noted only in hagfish⁷, guinea pig⁷ and casiragua⁶ insulins. These latter insulins lack the zinc-coordinating B₁₀ His together with several other residues on the dimer-dimer interface of bovine insulin which have been implicated in zinc binding and hexamer formation⁸. As these residues are conserved in porcupine insulin, an alternative explanation is required to account for its failure to bind zinc.

The CD spectrum of porcupine insulin shows concentration independence in both the near and far UV (Fig. 3). The intensity of the tyrosyl CD band of bovine insulin at 275 nm has been shown to be dependent on concentration and has been largely

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30												
A Chain																																										
Bovine	Gly	Ile	Val	Glu	Gln	Cys	Cys	Ala	Ser	Val	Cys	Ser	Leu	Tyr	Gln	Leu	Glu	Asn	Tyr	Cys	Asn																					
Porcupine	Asp				Thr				Gly		Gln																															
B Chain																																										
Bovine	Phe	Val	Asn	Gln	His	Leu	Cys	Gly	Ser	His	Leu	Val	Glu	Ala	Leu	Tyr	Leu	Val	Cys	Gly	Glu	Arg	Gly	Phe	Phe	Tyr	Thr	Pro	Lys	Ala												
Porcupine																					Asn				Asp																Arg	

Fig. 1 Comparison of the primary structures of bovine and porcupine.

assigned to perturbation of the tyrosines at B₂₆ and B₁₆, which become buried as the hormone dimerizes (W. Strassburger, J. Fleischauer and A.W., unpublished results). Because these residues are conserved in porcupine insulin, one would expect a concentration dependence at 275 nm if dimerization occurred. Therefore, the lack of concentration dependence is further proof of the inability of porcupine insulin to dimerize. As porcupine insulin is monomeric, its CD spectrum is compared with that of monomeric bovine insulin (that is, at concentrations below 10⁻⁶M). The CD spectra of the two insulins can be seen to be more similar at 275 nm (Fig. 4). However, the positive band at 196 nm in bovine insulin is blue-shifted to 193 nm in porcupine insulin and is reduced in ellipticity. The negative extremum found at 208 nm in bovine insulin also shows reduced ellipticity in porcupine insulin. The ellipticity of the bands at 196 and 208 nm are dependent on backbone conformation rather than on side-chain chromophores and have been largely assigned to α -helix. These data suggest that at pH 8.4 the tertiary structures of porcupine and bovine insulins are different. On acidification to pH 2, the CD spectrum of porcupine insulin more closely resembles that of bovine insulin measured at neutral pH; at 275 nm the two spectra are almost indistinguishable, although differences at 196 nm are still apparent (Fig. 4). The acidification of porcupine insulin seeks to alter its tertiary structure towards a more 'bovine-like' insulin conformation.

We explain these observations as follows. In bovine and most other insulins the conformation of the hormone may be stabilized by a complementarity of charge between B₂₂ Arg guanidinium group and A₂₁ Asn carboxylate⁸. Both enzymatic

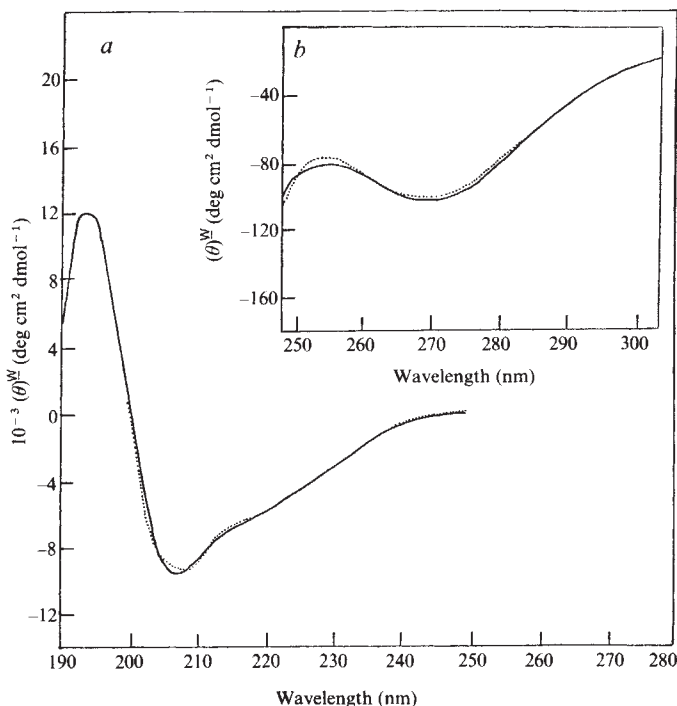


Fig. 3 Circular dichroism spectrum of porcupine insulin as a function of concentration. —, 4 × 10⁻⁵M porcupine insulin; ····, 4 × 10⁻⁶M porcupine insulin in 0.025 M Tris buffer, pH 8.4, zinc free. Bovine insulin has a concentration-dependent CD spectrum at both 222 nm and 275 nm (ref. 5).

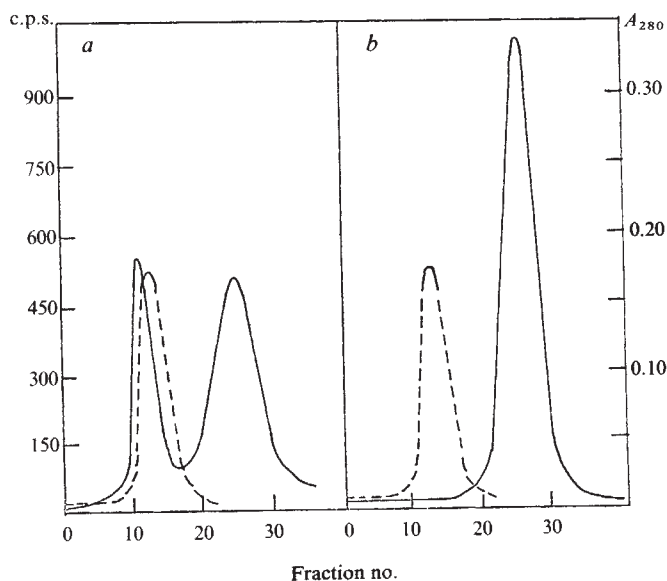


Fig. 2 The formation of ⁶⁵Zn-insulin complexes by bovine (a) and porcupine (b) insulins and their separation from free ⁶⁵Zn ions on a Sephadex G-25 column (1.5 × 30 cm) at pH 7.4 in a Veronal buffer⁴, 3 ml fraction volumes collected. ---, Insulin; —, ⁶⁵Zn.

cleavage of the A₂₁ Asn (ref. 9) and substitution of the B₂₂ Arg (ref. 10) result in a modification of the tertiary structure of the hormone and affect its ability to aggregate^{5,9,10}. The substitution B₂₂ Arg by Asp in porcupine insulin (Fig. 1) will bring two negative charges into close proximity unless there is a change of conformation, which may also result in a loss of zinc binding and self-association. At low pH, protonation of either B₂₂ Asp or the A₂₁ α -carboxylate may remove charge-charge repulsion and allow a close approach of the B₂₂-A₂₁ residues, resulting in the change of tertiary structure of porcupine insulin towards the conformation of the bovine molecule. A similar mechanism has been suggested for the effect of acidification of guinea pig insulin¹¹.

The affinity of binding of porcupine insulin to rat liver plasma membrane insulin receptors and its ability to stimulate the conversion of [3-³H]glucose into lipid by isolated rat fat cells were determined. The most notable property of porcupine insulin revealed by these assays, apart from the decrease in biological potency of the hormone, was the discrepancy between the receptor binding affinity, 25%, and the biological activity, 4%, of the hormone (both relative to bovine insulin). Similar discrepancies have been found in hagfish insulin¹² and in several

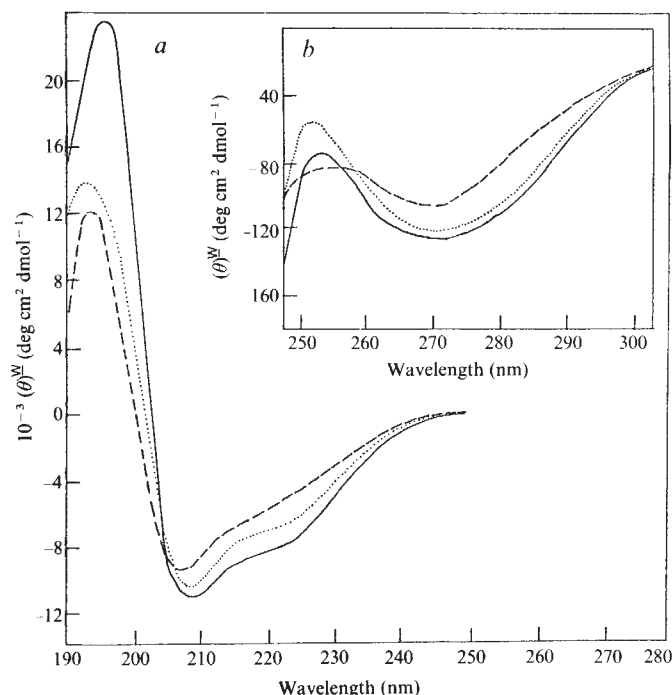


Fig. 4 Circular dichroism spectrum of porcine insulin at pH 2.0 and 8.4 and of bovine insulin at pH 7.8. ---, Porcine insulin 4×10^{-5} M, pH 8.4; ····, porcine insulin 4×10^{-5} M, pH 2.0; —, bovine insulin 4×10^{-6} M, pH 7.8. All are in 0.025 M Tris buffer at the pH indicated, zinc free.

insulin analogues¹³. We are unable to detect a common factor between these insulins and cannot explain the discrepancy. The decrease in potency of porcine insulin is probably due to the substitution of the B₂₂ Arg, because this is the only residue not conserved in the putative receptor binding site proposed by Pullen *et al.*¹⁴. A similar substitution, B₂₂ Arg to Asp, arises in guinea pig insulin and has been cited by Zimmerman *et al.*¹⁵ as being responsible for the low potency of this insulin. A recently synthesized analogue of bovine insulin substituted only at B₂₂ Arg to Asp has been reported to have a reduced potency (K. Yueh-ting, personal communication).

Other substitutions in the sequence of porcine insulin which might also contribute to its decrease in binding affinity are at A₄ Glu to Asp, B₂₁ Glu to Asn, and B₂₇ Thr to Arg, all of which lie on the periphery of the receptor binding region of the hormone¹⁴. Note also that hagfish insulin, which also has a substitution at B₂₁ Glu (Glu to Val), has a greatly reduced potency¹². This insulin has conserved all the residues involved in receptor binding and has a nearly identical conformation to that of bovine insulin¹⁶. Thus, it is likely that the substitution at B₂₁ makes an important contribution to the decrease in potency of porcine insulin.

Examination of the primary structure of four New World hystricomorph insulins reveals no one amino acid substitution that is characteristic of this suborder³. All the insulins share the property of low potency (blood glucose lowering, or [³-H]glucose conversion into lipid) when compared with bovine or porcine insulin. Scrutiny of the porcine insulin sequence also fails to show any characteristic amino acid substitutions. Low potency persists. We speculate that hystricomorph rodents, both New and Old World¹⁷, have low-potency insulins because of a common evolutionary pressure. This pressure has resulted in a reduction of the role of insulin as a modifier of glucose metabolism, as well as a change in the insulin receptor⁵. We speculate that another function of insulin has been consequently upgraded. One such function could be its mitogenic activity.

We thank Professor S. Gombe for organizing the collection of porcine pancreas, and Drs A. J. B. Haigh and D. G. Rae (Wellcome Kenya Ltd) for maintaining contact and successfully

preserving and despatching the glands. R.H. thanks the SRC for a CASE Studentship.

Received 28 February; accepted 9 June 1980.

1. Blundell, T. L. & Wood, S. P. *Nature* **257**, 197–203 (1975).
2. Smith, L. F. *Diabetes* **21**, Suppl. 2, 457–460 (1972).
3. Horuk, R. *et al.* *Nature* **279**, 439–440 (1979).
4. Zimmerman, A. E., Kells, D. I. L. & Yip, C. C. *Biochem. biophys. Res. Commun.* **46**, 2127–2133 (1972).
5. Wood, S. P., Blundell, T. L., Wollmer, A., Lazarus, N. R. & Neville, R. W. J. *Eur. J. Biochem.* **55**, 531–542 (1975).
6. Horuk, R. *et al.* in *Hormones and Cell Regulation* (eds Dumont, J. E. & Nunez, J.) (Elsevier, Amsterdam, in the press).
7. Peterson, J. D., Steiner, D. F., Emdin, S. O. & Falkmer, S. J. *biol. Chem.* **250**, 5183–5191 (1975).
8. Blundell, T. L., Hodgkin, D. C., Dodson, G. G. & Mercola, D. *Adv. Protein Chem.* **26**, 279–402 (1972).
9. Goldman, J. & Carpenter, F. H. *Biochemistry* **13**, 4566–4574 (1974).
10. Zimmerman, A. E. & Yip, C. C. *J. biol. Chem.* **249**, 4021–4025 (1974).
11. Wood, S. P. thesis, Univ. Sussex (1976).
12. Emdin, S. O., Gammeltoft, S. & Gliemann, J. *J. biol. Chem.* **252**, 602–608 (1977).
13. Sönksen, P. H. *Nature* **282**, 11–12 (1979).
14. Pullen, R. A. *et al.* *Nature* **259**, 369–373 (1976).
15. Zimmerman, A. E., Moule, M. L. & Yip, C. C. *J. biol. Chem.* **249**, 4026–4029 (1974).
16. Cutfield, J. F. *et al.* *J. molec. Biol.* **132**, 85–100 (1979).
17. Wood, A. E. in *The Biology of Hystricomorph Rodents* (eds Rowlands, I. W. & Weir, B. J.) 21–51 (Academic, London, 1974).

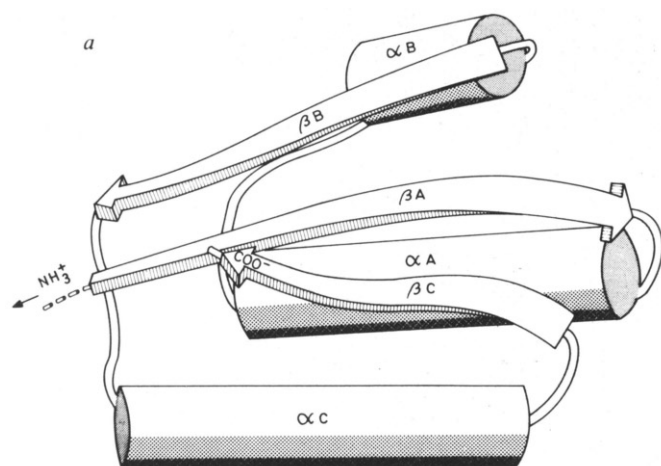
Crystal structure of a ribosomal component at 2.6 Å resolution

Marie Leijonmarck, Stig Eriksson & Anders Liljas

Institute of Molecular Biology, Uppsala University, Box 562, S-751 22 Uppsala, Sweden

Structural information about the ribosomes has been sought in many ways^{1,2}. It is known that the 50S subunit of *Escherichia coli* ribosomes contains four copies of the ribosomal protein L7/L12 (refs 3–7) in a unique region^{8,9}. A complex can be isolated consisting of those four copies bound to another ribosomal protein, L10 (refs 10–13). L7/L12 aggregates into elongated dimers in solution^{14,15}. Ribosomes deprived of L7/L12 have a reduced affinity for supernatant factors such as EF-Tu and EF-G and almost no capacity for factor-dependent GTP hydrolysis^{16–18}. The sequence of the 120 amino acid residues in L7/L12 from *E. coli* has been determined¹⁹, and two domains of L7/L12 (residues 1–36 and 53–120, respectively) have been crystallized separately²⁰. We now report the first crystal structure of a ribosomal component, a C-terminal fragment (CTF) of the protein L7/L12 from *E. coli*, determined to 2.6 Å resolution. The CTF has a compact, plum-shaped tertiary structure with a high content of secondary structure. The three α -helices and three β -strands in CTF account for 32% α -helix and 14% β -structure in the intact L7/L12 protein. A binding site for anions shows some resemblance to the known nucleotide binding site in many proteins.

Ribosomes, subunits, the protein L7/L12 and its fragments were obtained as described earlier^{11,20}. The crystallization of CTF started with a solution containing 0.2 M citrate pH 3, 10 mM MgSO₄, 0.8 M (NH₄)₂SO₄ and 2–10 mg ml⁻¹ protein. Crystals were obtained in hanging drops at 20 °C by vapour diffusion against 4 M (NH₄)₂SO₄. These bipyramidal crystals grew to a maximum length of about 0.4 mm over a period of 1–2 months. Before use the crystals were washed in 4 M (NH₄)₂SO₄. Native data were collected on a crystal in 4 M (NH₄)₂SO₄ at pH 4.2. Two heavy atom derivatives were used: 2 mM *p*-chloromercuribenzenesulphonic acid (PCMBs) at pH 5.2 and 1 mM KAu(CN)₂ at pH 4.2. They were soaked for 3 days. These crystals belong to the space group P4₃2₁2 with cell dimensions $a = b = 54.7$ Å and $c = 42.5$ Å and contain one molecule of CTF (molecular weight $\approx 6,900$) per asymmetric unit. X-ray intensities were collected at +4 °C on a Philips-Stoe computer-controlled 4-circle diffractometer²¹. The 2,218 independent reflections in the 2.6-Å resolution sphere were collected on one native crystal. For the heavy atom derivatives, Friedel pairs were measured at $\pm 2\theta$ in blocks of 100 reflections. Data were cor-



rected for Lorentz-polarization effects and for radiation damage and a semi-empirical absorption correction was applied²².

The heavy atom positions were determined from three-dimensional difference Patterson functions, calculated with coefficients including the anomalous scattering²³. The Patterson map of the PCMBs derivative was interpretable in terms of a single PCMBs site with the Hg atom as well as the S atom on the crystallographic 2-fold axis. The major site of the gold derivative was obvious from the Patterson map, whereas minor sites were derived from cross-phased difference Fourier maps. With phase angles in the two enantiomorphic space groups derived from the PCMBs derivative alone, two difference Fouriers of the gold cyanide derivatives were calculated. The clear difference in peak height showed the space group to be $P4_32_12$. The final state of the least squares refinement is shown in Table 1. Using the phase angles obtained, a 'best' native Fourier²⁴ was calculated.

The electron density map shows a clear distinction between solvent and protein. The protein part was interpreted as a single continuous polypeptide chain. A unique alignment could be made of the amino acid sequence (residues 53–120) with regard to the chain tracing. Rough coordinates were measured from the map as a guideline for the model building on a Vector General 3404 graphical display unit. The program package FRODO by Jones²⁵ was used.

The C-terminal fragment of L7/L12 is a plum-shaped molecule of approximate dimensions $20 \times 20 \times 35$ Å. It is composed of three α -helices and three β -strands, all lying roughly along the long axis of the molecule (Fig. 1). The three helices form one layer and the antiparallel β -strands form a second layer with the

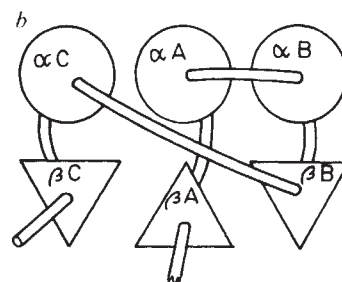


Fig. 1 Idealized representations of the L7/L12 C-terminal domain. *a*, Helices are represented by cylinders and β -strands as arrows pointing in the direction from N- to C-terminus. *b*, The structure seen from one edge. Δ , β -strand pointing into the paper; ∇ , β -strand pointing out of the paper; $\circ$, α -helix. This simple two-layer structure is composed of three helix-strand couples in the sequence $(\beta\alpha)_A(\alpha\beta)_B(\alpha\beta)_C$.

usual twist²⁶. We have found no other protein containing this simple arrangement of secondary structure elements.

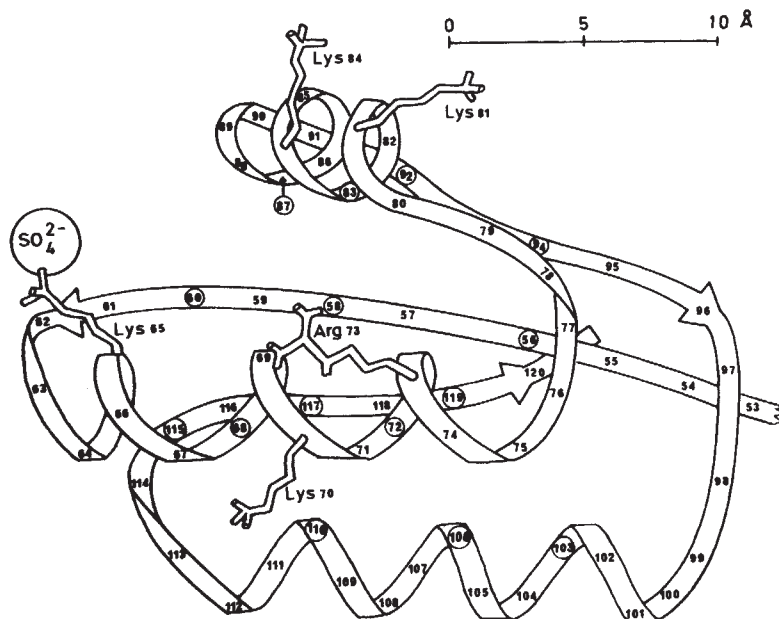
L7/L12 has a high proportion of charged residues. Nevertheless, all the residues between the two layers of secondary structure are hydrocarbon side chains (Fig. 2). These residues are never exchanged for polar residues in any of the three known sequences (Fig. 3).

A surprisingly large surface area on the β -sheet side of the CTF molecule is composed of only hydrophobic side chains (Ile 57, Ala 61, Leu 78, Ala 90, Pro 91 and Ala 93). This hydrophobic surface may ultimately be found to be a surface for contact of CTF with other regions of the L7/L12 polypeptide or for other proteins²⁷. The hydrophobic surface of CTF is lined by acidic residues, Asp 55, Glu 82, Asp 85, Glu 96 and Glu 118 and the C-terminal carboxyl group.

CTF obviously has a compact domain structure. This agrees well with the following observations: the spontaneous cleavage²⁰ seems to have occurred where the polypeptide enters into the domain; limited proteolysis experiments also yield a very similar fragment²⁷; a comparison of the L7/L12 sequences from three species shows that the region just before residue 53 is hydrophilic, highly variable and has frequent deletions⁵.

CTF has about 38 residues (56%) in α -helices and 17 residues (25%) in β -strands, 32% and 14%, respectively, of the 120 residues in L7/L12. A high helical content for L7/L12 has been observed by many investigators from circular dichroism (CD) spectra^{3,28–31}. Two groups have observed 10–20% β -structure^{3,31} by analysis of the entire far-UV region of the CD spectrum. With so much β -structure in CTF, there would be

Fig. 2 The structure of the L7/L12 C-terminal domain shown as a ribbon. The approximate location of each residue is indicated. This view of the molecule is opposite to the one in Fig. 1. The residues pointing into the hydrophobic core are in circles. This side of the molecule has five basic residues specially indicated here, without counter ions from the protein. A very significant electron density, best interpreted as a bound sulphate ion, is observed near Gly 62. This binding has some similarities to the nucleotide binding family of proteins.



BOOK REVIEWS

The growing pains of Robert Oppenheimer

Rudolf Peierls

Robert Oppenheimer: Letters and Recollections. Edited by Alice Kimball Smith and Charles Weiner. Pp.250. (Harvard University Press: 1980.) \$20, £12.

ROBERT Oppenheimer was one of the most famous scientists of his time, not so much for his contributions to theoretical physics, although these were substantial enough, but as mentor of a large number of successful theoreticians, and above all for his direction of the Los Alamos Laboratory at which the atom bomb was developed and produced during the Second World War.

The impressions he made on superficial acquaintance were very different at different times and on different occasions, and sometimes appeared hard to reconcile. The shy and introverted young boy and student became the successful director of a vast enterprise, successful very largely by his skill in dealing with people. The young researcher, whose professors had doubts about his scientific future because he had difficulty in expressing himself in writing, became famous for always finding the *mot juste* to sum up a situation. His lightning speed of understanding, which often produced the answer or a comment before his questioner had completed half his explanation, would on occasions make him miss the point of an important new idea, and he could meet presentations of such ideas by young colleagues with withering criticism. The ivory-tower scholar who did not know or care much about public affairs found his social conscience in the late 1930s by taking an interest in liberal and left-wing groups (an interest which formed part of the accusations in the security hearings against him), and later became a national figure in the debates about international relations and defence policy.

The present volume traces the development of this complex personality. The editors, Alice Kimball Smith and Charles Weiner are singularly well qualified for their task. Mrs. Smith was at Los Alamos from the start until the end of the War, and therefore well acquainted with the atmosphere of the place and with many of its members. She later made a study of the part played by scientists in the post-war period in influencing legislation towards

civilian control of atomic energy. Dr Weiner is a physicist concerned with recent scientific history, who has recorded interviews with many distinguished physicists.

In describing themselves as editors of Oppenheimer's letters and recollections, they do not really do justice to their work. The letters are supplemented, where there are gaps, not only by passages of Oppenheimer's recollections taken from later interviews, but also by comments from friends, colleagues and pupils, painstakingly collected and sensitively evaluated. In many places the editors' comments, based on this material, are more informative than the letters. They not only identify many of the people and events referred to in the letters, but also summarize events and developments not mentioned in the letters or which took place in periods from which no letters have survived. The book is therefore really intermediate between edited correspondence and a serious biography.

The young Oppenheimer was not the type to keep copies of his letters, and not many of the recipients preserved them. The fact that there remain enough letters from the pre-War period to provide a clear picture of Oppenheimer's attitudes and development during that time is largely due to the few friends who kept the letters and made them available. One of these is Herbert W. Smith, his English teacher in the Ethical Culture School; letters to him are most numerous of any of the correspondents. In the earlier part Oppenheimer sends him essays and verse; his evident desire for approval makes Smith a kind of father figure.

Almost as numerous are the letters to Oppenheimer's brother, Frank, in which the elder brother gives much advice and exhortation, and shows great warmth and affection, though sounding a little patronizing at times. Amongst other sizeable collections of letters are those to Francis Ferguson, whom he first met in his final year at the Ethical Culture School and with whom a close and sometimes tempestuous friendship developed. Of a later period, the physicist George Uhlenbeck, who was a close friend, preserved many interesting letters.

The first three chapters are based on these letters and a few others,

IMAGE
UNAVAILABLE
FOR
COPYRIGHT
REASONS

Robert Oppenheimer, 1926 or 1927 — "the young researcher whose professors had doubts about his scientific future . . .".

IMAGE
UNAVAILABLE
FOR
COPYRIGHT
REASONS

Oppenheimer in later years, after his success as director of Los Alamos and the ordeal of the "Oppenheimer Trials".

supplemented by other information collected by the editors. From his home background we follow Robert to the Ethical Culture School, where he was an excellent scholar in all subjects, and where his interest in science was first aroused through mineralogy. After leaving school and after a year's convalescence from dysentery contracted during a trip to Europe, he visited New Mexico in the summer with his teacher, Herbert Smith. This was the beginning of his deep attachment to the south-west. In 1922, aged 17, he entered Harvard. He spread his studies over a wide range of subjects, mostly on the literary side but including chemistry, in which he graduated in 1925, and some physics. The number and diversity of courses, and the amount of reading he managed, must have meant hard work; at the same time he learned to make contact with people and form friendships.

In his final year at Harvard an important influence was Percy Bridgman, in whose laboratory Oppenheimer did some experimental work, and his decision to go for postgraduate study in physics probably owed much to Bridgman. He decided to go to Cambridge. His studies there did not work out too well. He was not accepted to work under Rutherford, as he had hoped, but did some experimental research under J.J. Thomson, which did not make much progress. However, he had the chance at Cambridge of meeting many distinguished people, including Niels Bohr and other theoreticians, and it became clear to him then that he wanted to do theoretical work. During the Cambridge period he had times of very deep depression and unhappiness, but he was able to deal with his own disturbed state more effectively than any doctors he consulted.

After a year at Cambridge he moved to Göttingen, a congenial environment, to work with Max Born. This was the exciting time when quantum mechanics emerged, and the beginning of Oppenheimer's passionate devotion to theoretical physics, and also of his constructive contributions to the subject. He qualified in one year for a PhD, and returned to the United States in 1927. After a semester at Harvard and one at the California Institute of Technology, he spent another year in Europe, visiting Ehrenfest, Kramers and Pauli. Then he moved to California, holding appointments at Caltech and Berkeley, dividing his time between the two places. He liked this arrangement because at Caltech there were many theoreticians from whom he could learn, while at Berkeley he was practically the only modern theoretical physicist, and thus could contribute more to teaching and research. This unusual arrangement continued until he became involved with atomic energy in 1942. It was here that he started to gather a group of graduate students, who usually followed his annual migration — and that he developed his

power as a teacher and leader in research. It was in the California years that he first became concerned with the world around him and with politics. It was in that period that he got married.

He was now an established and respected member of the physics community, as expressed in his election to the National Academy in 1941.

In 1942, when work on the atomic bomb was becoming a serious project, he was asked to take charge of work on fast-neutron fission, which was related to the problems of design and performance of a bomb. The period from this time until the end of the War is dealt with in Chapter 4. Here the task of the editors changes. It is no longer a question of searching out the few surviving letters and filling gaps, but rather to select, from amongst the voluminous files of official correspondence, the letters which are most informative of Oppenheimer's attitude and ways.

In coordinating the experimental and theoretical work carried out in different universities, he immediately showed his characteristic ability for summarizing, for stating the main issues clearly and for formulating questions. It soon became clear that efficiency demanded that this work be concentrated in a single centre, and he became involved in the planning of the new laboratory and the selection of a site. At some stage General Groves, head of the 'Manhattan District' (the code name for the atomic weapons project), decided to appoint Oppenheimer the director of the new laboratory. This was a daring decision, considering that Oppenheimer was a theoretician, not experienced in administration and had the reputation of being rather disorganized in practical matters. Yet the decision was amply justified by its success. Chapter 4 helps to understand why. In addition to his lucidity in recognizing and expressing the main issues, Oppenheimer managed to deal with people in a manner which made them feel that they were respected, and gave them the confidence that their views and their needs

were taken into account. The letters give examples of the many large and small problems that the director had to handle, and the pressure under which he was working. One of the surprising facts is how much mutual respect, and indeed affection, developed between Oppenheimer and Groves, in spite of the enormous disparity between their personalities.

Chapter 5 covers the three months from the end of the War to Oppenheimer's departure from Los Alamos. The letters are now concerned with post-War plans for Oppenheimer, positions for the staff of Los Alamos and the organization of atomic-energy research, and reflect the burden of responsibility that scientists carried because of their work on the bomb. The last document by Oppenheimer is his speech at Los Alamos, in which he discusses the future of international relations and the idea of international control of atomic energy.

The editors decided, wisely, to terminate the collection of letters at this point, since from then on the glare of publicity which fell on him, the great volume of correspondence and his membership of numerous official committees, changed the style of his letters, which tended to become more official and more guarded. They do not give the kind of insight which makes the letters reproduced in the book such fascinating reading. The book therefore does not, except in a brief résumé, deal with the post-Los Alamos Oppenheimer, with his directorship of the Institute for Advanced Study in Princeton or with the security hearings usually called the "Oppenheimer Trials", but there are other sources to cover these matters. However, the period it does cover, the growth of an unusually gifted and sensitive person to maturity, and to his success in a gruelling job, is of absorbing interest. □

Sir Rudolf Peierls is Emeritus Wykeham Professor of Physics at the University of Oxford. He spent 18 months at Los Alamos working under Oppenheimer.

Pluto, the unique planet

J. Veverka

The Planet Pluto. By A. J. Whyte. Pp. 147. (Pergamon: 1980.) \$17.50, £8.75.

FIFTY years have passed since the discovery of Pluto on February 18, 1930, but until recently the planet's physical properties were still subject to much debate. The discovery of a satellite by Christy and Harrington in 1978 made it possible to determine the mass of the planet, and recent speckle interferometry by Arnold, Boksenberg and Sargent produced the first

good measurement of the radius. Finally we can say something about Pluto's mean density and composition.

Within half a century our conception of Pluto has changed from that of a rocky body, about the size of Earth, to an icy double-planet so tiny that it is dwarfed by such satellites as Ganymede and Titan. Latest data indicate that the planet is only some 1,500 km in radius (slightly smaller than our Moon, but three times bigger than the largest asteroid). Its companion, Charon, a relative giant, is almost half the size of the primary (our Moon is only a quarter as big as Earth). The average density of Pluto is probably less than 1.5 g cm⁻³, indicating a predominantly icy constitution. It is still uncertain whether

Pluto does or does not have an atmosphere, but methane frost on its surface was discovered in 1976 by Cruikshank, Pilcher and Morrison.

Whyte's *The Planet Pluto* is one of two very timely recent books about this unique world. Both can be recommended highly. The second book is W. G. Hoyt's *Planets X and Pluto* (University of Arizona Press), a well-documented, readable account of the discovery and subsequent studies of the planet, based largely on Lowell Observatory records.

Whyte's book is a succinct, thorough summary of our growing knowledge of Pluto. Results are given more or less chronologically without critical comment. This type of presentation has its advantages, but there are occasions where one wishes the author had allowed himself to be more critical. Obvious examples involve Kiladze's numerology, Brady's planet, certain very noisy spectra over-interpreted in fantastic fashions and pedantic models of the planet's interior calculated before the mass or size of Pluto were known.

There are only occasional typos and the odd trivial error, and few important references are omitted. Most notably, no mention is made of such important papers as those of Seidemann *et al.* (*Publ*

astr. Soc. Pacif. **84**, 858) and Goldreich and Ward (*Publ. astr. Soc. Pacif.* **84**, 737) which appeared in 1972 and demonstrated that Brady's huge trans-Neptunian planet is an impossibility. The references are up to date until July 1979, and thus do not include some important new results such as the speckle determination of the radius of Pluto by Arnold and his colleagues.

In spite of its high price, the book is recommended heartily to anyone interested in Pluto. It is hoped that it will make a contribution to a much-needed re-evaluation of Pluto's importance to our understanding of the Solar System. Perhaps more astronomers will see fit to observe the planet with the techniques now available while it is near perihelion during the next two decades. Pluto has emerged as a unique body in our Solar System and as such is worthy of closer study, and especially of direct exploration by spacecraft. Opportunities exist to fly-by Jupiter — and have another look at Io — and arrive at Pluto before the end of the century. □

J. Veverka is at the Laboratory for Planetary Studies, Cornell University, New York.

End of a chapter in population genetics

Alan Robertson

Mathematical Population Genetics. Biomathematics Vol. 9. By W.J. Ewens. Pp.322. (Springer-Verlag: 1980.) DM 59, \$33.10.

SINCE the rediscovery of Mendelian segregation, the combinatorial properties of the phenomenon have been fascinating to the mathematician (as indeed they were to Mendel himself). What Ewens calls the "golden age" of mathematical population genetics was dominated by Fisher, Haldane and Wright, the first two with professional training and the latter an inspired amateur. For a time, the theory was far ahead of the relevant evidence but the emergence of new experimental techniques (gel electrophoresis, amino-acid sequencing of proteins and, more recently, DNA sequencing) has not only provided a large body of data, though mostly concerned with protein evolution, but has also stimulated new conceptual approaches, such as the misnamed 'non-Darwinian evolution', and a great deal of sophisticated mathematics. For instance, our view of the basis of mutation at individual loci is quite altered — we think no longer of 'forward' and 'back' mutation but of an infinite progression which almost never retraces its path and

which has led to an 'infinite allele' theory of some elegance.

Many recent theoretical papers seem to have been written for mathematicians rather than biologists. Ewens does an excellent job of presenting the different contributions, many from himself or his students. Aimed at the graduate or research student, two threads are neatly interwoven throughout. Having first presented the theoretical approach to a topic, he ends with his own view of its present biological status. The book deals mainly with the statistics of genes in populations and less than 15 pages are concerned with quantitative genetics. Its value to the reader obviously depends on the latter's mathematical level. I have certainly found it useful and stimulating — an indication not only of my mathematics but that I share the author's prejudices.

Perhaps, too, this volume marks the end of a chapter — of a theory of evolution in terms of the substitution of single mutations. Papers dealing with the mathematical properties of systems of tandemly repeated loci have already appeared (though none are referred to here), and current issues of *Cell* and *PNAS* are full of discussions of transposable sequences, a new kind of genetic polymorphism, arising from a novel source of variation. □

Alan Robertson is Professor at the Institute of Genetics, University of Edinburgh.

The potential of peat

Hamish Anderson

Peat: Industrial Chemistry and Technology. By Charles H. Fuchsman. Pp.279. (Academic: 1980.) \$28, £15.80.

AT A time when many countries, faced with ever-declining supplies and resources of basic raw materials, are currently re-assessing the potential of alternative sources of energy and of feedstocks for the chemical industry, the appearance of Dr Fuchsman's book is particularly welcome and opportune. Indeed, this publication constitutes the first real attempt in the English language to collect and collate information on the chemistry and technology of an industry which, until now, has centred largely in eastern and northern Europe. Based on data judiciously selected from over 450 references, including many from Soviet publications not readily accessible to Western scientists, the author presents a most readable and concise review of the nature of organic components in peat and of the alternative technologies available for converting these components into commercially useful and value-added products.

Following a general introduction on the nature and distribution of peat resources and the history of peat utilization for industrial purposes, Chapter 2 presents a somewhat superficial account of the classification and chemical characterization of different bog and peat types, as a result of which the complicated and often ambiguous terminology and nomenclature may trouble the non-specialist. It is also surprising that sapropel, a material now widely used in the USSR, is considered to be of little importance to chemical technology.

The extraction of peat with organic solvents is dealt with in considerable detail in Chapters 3–5, which help to clarify a murky area of the peat research effort. The chemistry and utilization of peat carbohydrates and hydrolysates of peat polysaccharides are comprehensively covered in Chapters 6–8, followed by an account of the hydrolytic production of furfural, lactic and glycolic acids and other sugar by-products. A further use of these hydrolysates is the indirect production of proteins and fats by the culture of yeasts — a process which has a long and honourable

The second edition of the *Glossary of Geology*, edited by Robert L. Bates and Julia A. Jackson, has been published by the American Geological Institute, Falls Church, Vermont. The updated *Glossary* includes 36,000 terms, compared to 33,000 in the 1972 edition, and is available from the AGI, price \$60.

history in peat technology.

In a book which deals largely with the subject from an industrial viewpoint, it is surprising to find that the longest section (Chapter 12) is devoted to a review of the occurrence and structural chemistry of humic acids and lignin. By contrast, Chapters 13–15 provide a most valuable account of the theory and practice of peat carbonization, specifications for coke and related products, and the production and utilization of distillates and gases from peat pyrolysis.

Chapter 16 on chemical methods of peat

analysis is perhaps the weakest in the entire book. In a work containing such an extensive and well-documented bibliography up to, and including, 1978, it is surprising to be presented with many analytical methods which are out of date. Thus much stress is placed on the use of paper chromatography for the estimation of components of hydrolysates, disregarding modern liquid and gas chromatographic techniques for the analysis of waxes, steroids, terpenes, phenols, sugars and amino acids. Automated flow analysis methods are not mentioned. The concluding three chapters

deal with the scale of peat processing operations, economic and social implications, environmental impact and pollution control.

Overall, Charles Fuchsman's book fills an important gap in existing literature and will be widely welcomed not only by specialists in the field concerned but by all those engaged in peat research and development. □

Hamish Anderson is a Principal Scientific Officer at the Macaulay Institute for Soil Research, Aberdeen.

Art of the insoluble

R. H. Pain

Hydrophobic Interactions. By A. Ben-Naim. Pp.311. (Plenum: 1980.) \$32.50, £20.48.

MEDAWAR has reminded us that the pursuit of science is the art of the soluble. This book is about the art of the insoluble.

To judge by the number of monographs recently published on the subject, interest in the hydrophobic effect and its involvement in the stabilization of biological structures is at a high point.

Kauzmann in 1959 (*Adv. Protein Chem.*

14, 1) drew attention to the fact that the association of non-polar groups in an aqueous medium exhibits thermodynamic properties which can only be explained in terms of changes in the entropy of the solvent. Such groups, which are present on many biological molecules, come together not so much due to mutual attraction as to a mutual distaste for water. These concepts have been developed by several groups to provide a valuable tool with which biochemists can rationalize a wide range of interactions. What may be called the thermodynamic approach to hydrophobic interactions is widely accepted and remains useful because of the still limited understanding of the structure of water. The beauty of thermodynamics in such a situation lies in its ability to handle phenomena quantitatively without having to make any assumptions as to mechanism, and some of the workers in this field have vigorously eschewed any attempt to taint the purity of the approach by considering molecular mechanisms. The frustration which this kind of treatment apparently induces in biochemists is well illustrated by widespread references in the literature to 'water structures' and 'icebergs', and indeed there is much that we want and need to know about the details of the hydrophobic interaction.

The monograph by Ben-Naim attempts to steer a course somewhere in between the thermodynamic and the molecular-structure approaches. It is not in any sense a review, as the limited and selective list of references indicates, but an account of the author's own attempts over a period of years to tackle the problem of hydrophobicity. The first impression is one of clarity in his treatment of the subject. He leads the reader patiently, if at times paternally, through the development of his argument in a way which is refreshing and helpful. The argument is clarified by a series of diagrams which enable the model-orientated reader to grasp the significance of the equations. Frequent "Comment" paragraphs provide summary, objective assessment of the previous sections. It is, surprisingly, these comments which first make for a sense of unease, for the general impression conveyed is that more experi-

mental data are required to test the theory or that the theory being developed does not relate to real and highly complex macromolecular systems.

The author begins by drawing a useful distinction between the *direct* parts of the interaction between two particles in solution and the *indirect* part which comprises the hydrophobic interaction. His approach is based on that of statistical mechanics and he sets out to describe the interaction in terms of association rather than solubility. The contrasts with the thermodynamic approach are marked, in that the familiar description of the hydrophobic interaction in terms of characteristic changes in entropy and heat capacity is absent. In fact, consideration of these does not appear until three-quarters of the way through the book. The treatment progresses through discussion of the interaction between two simple particles to interaction between many solute particles, with particular emphasis on the non-additivity or cooperative aspects. The final chapter concerns the temperature and pressure dependence of hydrophobic interactions. Statistical mechanical derivations are assigned to appendices in order to make the text more readable. Various kinds of experiments are described, although some of these, such as those involving the association of carboxylic acids, are notoriously difficult to interpret in terms of the hydrophobic contribution.

This book is unlikely to be of value to the biochemist. Little light is shed on the four fundamental questions raised in the preface as to the nature and role of hydrophobic interactions — these are still best answered by intuitions and prejudices based largely on the thermodynamic approach. The book is a useful contribution for the hydrophobic buff, summarizing the application of the statistical mechanical approach. For my money, though, further understanding of the hydrophobic interaction will only come as we learn more about the behaviour of the solvent at the molecular level. □

R. H. Pain is Professor of Physical Biochemistry at the University of Newcastle upon Tyne.

DON'T MISS THE Grand Tour!

Last year at Jupiter. They are already nearing Saturn. Then it's on to Uranus, and Neptune, and possibly Pluto. Voyager photos are amazing, data from onboard experiments revolutionary. And full coverage of the tour can now be yours, along with features on nearly every other body in this solar system. So keep up with science - by subscribing to Planetary Astronomy.

Published bi-monthly by the James-Mims Observatory, Inc., a non-profit corporation. Subscriptions are \$5.00/yr., \$9.50/2 yrs (US), \$7.00/yr., \$13.50/2 yrs (Canada, Mexico) \$8.50/yr., \$16.00/2 yrs Other countries)

PLANETARY ASTRONOMY
P.O. Box 15854 Dept. N
Baton Rouge, LA 70895

Circle No. 04 on Reader Enquiry Card.